Gruther

proceedings of the

sixth annual conference on manual control

7-9 april 1970
wright patterson afb, ohio

air force institute of technology
air force flight dynamics laboratory

3624

20110106 180

this document has been approved for public release and sale; its distribution is unlimited.

proceedings of the

sixth annual conference on manual control

air force institute of technology
air force flight dynamics laboratory
wright patterson afb, ohio

7-9 april 1970

FOREWORD

This volume contains the proceedings of the Sixth Annual Conference on Manual Control held from 7 to 9 April, 1970, at the Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio. The program was divided into the following sessions: displays, neuro-muscular system analysis, man-machine analysis, manipulators and manipulation, advanced modeling techniques, describing function models, and describing function measurement techniques. Both formal and informal presentations were made; all of the formal and several of the informal papers are included in these two volumes. As indicated in the title, this was the 6th conference on manual control. The title, location, date, and report number for the first through the fifth conference proceedings are listed below for cross reference:

First Annual NASA-University Conference on Manual Control, University of Michigan, December 1964. (Proceedings not printed.)

Second Annual NASA-University Conference on Manual Control, MIT, Feb 28 to March 2, 1966, NASA SP-128.

Third Annual NASA-University Conference on Manual Control, University of Southern California, March 1-3, 1967, NASA SP-144.

Fourth Annual NASA-University Conference on Manual Control, University of Michigan, March 21-23, 1968, NASA SP-192.

Fifth Annual NASA-University Conference on Manual Control, MIT, March 27-29, 1969, NASA SP-215.

CONTENTS

page

Foreword iii
Summary ix

SESSION I. DISPLAY

1. The Use of Predictor Displays for Terminal Area Air Traffic Control 1
   William B. Rouse
2. An Investigation of Head Tracking Performance Using a Helmet Mounted Sight and Display System 7
   Allen H. Henke
3. Control Gains in Head-up Presentation 19
   J. M. Naish
4. Pilot Performance with a Simulated Pictorial Landing Display Including Different Conditions of Resolution And Update Rate 47
   T. Wempe and E. Palmer
5. Measurement and Analysis of Pilot Scanning Behavior During Simulated Instrument Approaches 83
   D. H. Weir and R. H. Klein
6. Tactical Utility Helicopter Information Transfer Study 109
   John A. Barnes
7. A Perspective Glideslope Indicating System 117
   Noel A. J. Van Houtte
8. An Integrated Display for Trajectory Control of Helicopters and VTOL Aircraft 133
   Theodor A. Dukes
9. The Statistical Phase Plane Display: A New Method for Monitoring Time Series 147
   Joseph G. Wohl

SESSION II. NEUROMUSCULAR SYSTEM ANALYSIS

10. Progress in Quantification of Muscular Action 165
    Julia T. Apter and William W. Graessley
11. A Model of Muscle Activity 177
    Gerald L. Gottlieb and Gyan C. Agarwal
12. On the Grading of Tension During Voluntary Contraction of Skeletal Muscle 195
    L. Ostroy, A. V. Phatak, and G. A. Bekey
13. The Dynamic Characteristics of Human Skeletal Muscle Modelled From Surface Stimulation 211
    J. A. Tennant
14. Sensory Feedback in Human Posture Control 249
    Lewis M. Nashner and Jacob L. Meiry

15. Experimental and Gerneralized Mathematical Analysis of the Human Motor Control System 269
Gerhard Vossius and Jurgen Werner
16. Direct Evidence Against Sampling at the Alpha Motoneuron .. 277
Gyan C. Agarwal and Gerald L. Gottlieb

SESSION III. MAN-VEHICLE ANALYSIS

17. Application of Parameter Optimization to Stability Augmentation Design 293
R. L. Stapleford
18. A Pilot-Vehicle System Approach to Longitudinal Flight Director Design 313
R. H. Klein, D. T. McRuer, and D. H. Weir
19. A Study of Relationships Between Aircraft System Performance and Pilot Ratings 339
W. C. Schultz, F. D. Newell and R. F. Whitbeck
20. "Paper Pilot" -- An Application of Pilot Models to Predict VTOL Flying Qualities in Precision Hover 349
R. O. Anderson and J. D. Dillow
21. Man/Machine Function Allocation of Flight Control 365
O. H. Lindquist
22. The Application of Human Operator Describing Function Theory to the Prediction of Tracking Performance in the Cheyenne Swiveling Gunner's Station 405
P. Briggs and L. G. Hofmann
23. On the Influence of Drugs on the Behavior of a Bicycle Rider 419
A. van Lunteren and H. G. Stassen

SESSION IV. MANIPULATORS AND MANIPULATION

24. Independent Changes in Visual and Kinaesthetic Feedback ... 439
Leslie Buck
25. Training for the Production of Memorized Movement Patterns 449
Terry R. Armstrong
26. On Simulated and Full-Scale Car Following 457
Robert E. Fenton, James H. Ott, and Ronald G. Rule
27. In-Flight Evaluation of Pilot's Controllers 479
David E. Frearson and James A. Townsend
28. A Computer Based Learning System for Inhibitory Teleoperator Control 481
Amos Freedy, Frederick Hull and John Lyman
29. On the Possibilities of Tactile Information Transmission For the Use in Arm Prostheses 513
H. G. Stassen, A. W. A. Meyer and A. van Lunteren
30. Can Proprioceptive Cues Unload the Human Operator? 535
Karl - Friedrich Kraiss

SESSION V. ADVANCED MODELING TECHNIQUES

31. Design of Corrective Dynamics Controllers 549
Ralph Mekel and Patrick Peruo, Jr.
32. A Two Dimensional, Pattern Recognizing, Adaptive Model of a Human Controller 553
D. W. Gilstad and K. S. Fu
33. An Input Adaptive, Pursuit Tracking Model of the Human Operator 571
John R. Ware
34. A Discrete Stochastic Optimal Control Model of the Human Operator 577
H. M. Paskin
35. A Model for Task Interference 585
William H. Levison
36. On How Often the Supervisor Should Sample 617
Thomas B. Sheridan
37. Systems Engineering, Social Science and the Quality of Urban Life 637
Ezra S. Krendel

SESSION VI. DESCRIBING FUNCTION MODELS

38. The Effects of Changes in Input Power Spectra on Human Operator Compensatory Tracking 677
John C. Heifferon and Russell A. Hannen
39. On the Variance of the Bicycle Rider's Behavior 701
A. van Lunteren and H. G. Stassen
40. On the Optimality of the Human Operator 723
Glenn A. Jackson
41. Visual Sources of Controller Remnant 737
William H. Levison
42. Research on a New Human Dynamic Response Test Battery 743
H. R. Jex and R. W. Allen
43. Identification of Pilot Dynamics with and without Motion Cues 779
E. P. Salmon and J. T. Gallagher

SESSION VII. DESCRIBING FUNCTION MEASUREMENT TECHNIQUES

44. A Simple Fourier Analysis Technique for Measuring the Dynamic Response of Manual Control Systems 785
R. Wade Allen and Henry R. Jex
45. A Comparison of Techniques for Measuring Human Operator Frequency Response 803
Richard S. Shirley
46. A Look at Pilot Modeling Techniques at Low Frequencies 871
Lawrence W. Taylor, Jr.

SUMMARY

Paul Pietrzak & Ron Anderson
Air Force Flight Dynamics Laboratory

Russell A. Hannen
Air Force Institute of Technology

The Sixth Annual Conference on Manual Control was held at the Air Force Institute of Technology from 7 to 9 April 1970. One hundred and fifteen specialists from the United States, Canada, Netherlands and Germany attended the conference, which was organized by the authors of this summary.

The formal program was divided into seven sessions. Highlights of the sessions are presented in the following paragraphs.

I. DISPLAY

The opening session on display involved a large number of informal presentations which covered a broad range of display situations and data gathering techniques. W. B. Rouse presented a plan to use a computer to aid the air traffic controller with a predictor display. The results of extensive flight and simulator tests of a head-up display system was presented by J. M. Naish. It is shown that a manual approach with HUD can allow the pilot greater command of the relevant information than is possible in automatic flight. T. Wempe and E. Palmer discussed the results of a study in which the resolution and update interval of a digitally simulated pictorial runway display were varied. The display had good pilot acceptance and could be degraded by large variations in resolution and update interval without affecting pilot performance.

Experimental measurements of pilot scanning and control response were reported by D. H. Weir and R. H. Klein. A new concept of measuring eye-point-of-regard was utilized in a simulated instrument approach of a subsonic jet transport. This data base when combined with concurrently measured pilot response data, will serve to validate and extend the theory of manual control displays. J. A. Barnes presented instrument fixation and scanning patterns obtained in flight for a variety of helicopter mission tasks.

Three presentations were involved with novel display systems for V/STOL vehicles. N.A.J. Van Houtte presented a glideslope indicating system generated on a large screen CRT using an analog line drawing scheme. Integration of all necessary flight information for precision trajectory control of helicopters into a single display was reported by T.A. Dukes. The basic invariant display configuration can be augmented to apply to any precision flying task and is adaptable to other VTOL aircraft. L.R. Young and G. Kemp discussed a "plan-view" VTOL display in which pitch and roll information are portrayed as projections of the aircraft Y-Z and X-Z planes respectively onto the ground. In addition prediction information is displayed in altitude and position.

R. E. Curry presented status reports of several efforts concerned with research and development of collision warning displays for light aircraft. Experiments in audio warning systems and a heads up display were discussed.

II. NEUROMUSCULAR SYSTEM ANALYSIS

The papers presented at this session showed the continuing interest in the analysis and modeling of muscles and muscle groups as well as the mathematical representation of the "fine structure" of the human operator while tracking. Four papers were addressed to the quantification of muscular behavior with a range of viewpoints and methods of attack. J. T. Apter and W. Graessley gave a progress report on the model they presented at last year's conference. A major step has been the development of a physical basis for the dependence of model parameters on the configuration of macromolecules of muscle to replace the empirical relation previously used. G.L. Gottlieb and G.C. Agarwal describe an analog matching of foot torque response to recorded EMG's. Still unresolved is the question of whether the EMG signal is a true measure of motoneuron activity or some function involving considerable dynamic processing of its own. If the former, it then represents the muscles controlling signal and the relationship between EMG and force describes the fundamental dynamics of muscle activity under normal operating activity.

The relationship of EMG to gradation of tension during voluntary contraction of skeletal muscle was investigated by L. Ostroy, A.V. Phatak and G.A. Bekey. Specifically the role of recruitment order of motor units which yield a linear EMG-tension relationship was investigated. J.A. Tennant presented a carefully detailed study in which he modeled the dynamic characteristics of a specific muscle group taking into account the changes in muscle tension due to the inertia of the moving masses. A multiloop control model for human posture control was synthesized in a paper by L. Nashner and J. Meiry. The model describes postural regulation about the ankle joints for very small disturbances during quiet standing, and vestibular regulation in the absence of visual and kinesthetic feedback cues.

G. Vossius and J. Werner presented models for human tracking movements in response to continuous and sampled input-signals. Agarwal and Gottlieb presented results of a physiological investigation of the sampling hypothesis at the alpha motoneuron in normal human subjects. The results indicate that the sampling hypothesis as proposed by Navas and Stark is not valid and the sampling behavior in the human motor system must be of central origin.

III. MAN-VEHICLE ANALYSIS

A wide variety of control situations was investigated in the papers included in this section. Applications included design of augmented flight control systems and flight director displays, prediction of pilot ratings of

handling qualities, man-machine function allocation, and the effect of drugs on the behavior of a bicycle rider.

A systematic procedure for the design of aircraft stability augmentation systems was proposed by R.L. Stapleford. The optimization problem was structured to include both pilot model and SAS feedbacks. An example application to the F-4 longitudinal axis demonstrated the feasibility of the approach. R. H. Klein, D.T. McRuer and D.H. Weir investigated the requirements and principles for analytical design of flight directors. Application of the theory of manual control displays produced the result that there are effective director/vehicle controlled element dynamics which are preferred from the standpoint of pilot response and system performance.

A series of three papers examined the relationship between man-machine systems performance and pilot ratings. The first study by W.C. Schultz, F.D. Newell and R.F. Whitbeck obtained pilot rating data for an ILS approach task. Analysis of the data resulted in no correlation between pilot rating and rms glide slope error. An investigation by Adams and Hatch presents a correlation between closed-loop pilot-aircraft system characteristics and pilot ratings previously obtained in several flight and simulator studies. The system characteristics are determined with pilot describing function models in closed-loop man-vehicle analysis. The third paper presented by R.O. Anderson and J.D. Dillow suggests a correlation of pilot model parameters and closed loop performance with the pilot opinion of VTOL hover flying qualities. The hypothesis that the pilot adjusts his method of control to minimize pilot ratings was investigated by implementing a digital computer program dubbed the 'paper pilot'. This ambitious program utilizes vehicle characteristics as input and computes pilot model parameters, closed loop performance, and pilot rating for the VTOL hover task in the longitudinal axis.

P. Briggs and L.G. Hofmann discussed the prediction of tracking performance for a helicopter swiveling gunner's station. The investigation resulted in the modification of existing describing function models for the accurate prediction of tracking performance.

Approaches to man/machine function allocation for flight crew stations were presented by O.H. Lindquist. It is suggested that system design should provide reserve crew "time capacity" to permit the man to meet the unprogramable contingencies of operation.

A. Van Lunteren and H. Stassen determined the influence of some drugs on the behavior of a rider stabilizing a bicycle simulator. The study was performed by measuring the drug effect on the parameters of the describing functions between the input of the rider, the frame angle, and the outputs, handle bar rotation and upper body angle.

IV. MANIPULATORS AND MANIPULATION

This session contained papers on a potpourri of topics according to the chairman. A quick reading of the titles confirms this statement as papers were presented on simulation, hardware development, remote manipulation, and other subjects which are not specifically manipulators.

L. Buck reported on experiments on manipulator-display compatability which attempted to localize feedback effects by varying kinesthetic and non-kinesthetic feedback independently. Results show degradation is similar for both kinesthetic feedback and visual feedback but recovery is more rapid from the former. The effect of augmented proprioceptive feedback information during training for the production of a memorized movement pattern was investigated by T.R. Armstrong.

R.E. Fenton, J.H. Ott and R.G. Rule described the components of a recently constructed moving base automobile simulator. Describing function data obtained in the simulator compare favorably with actual vehicle results. The program to evaluate four primary aircraft controllers was reported by J. Townsend and D. Drearson. The in-flight evaluation was conducted on a variable stability airplane for both up-and-away and approach tasks.

L.R. Young and P. Noggle discussed results of a controller which sensed the operator's force but moved in response to an external electrical signal giving the operator kinesthetic cues which could reduce his delay time or lead. L. Young then described a "yaw chair" that has been implemented to respond to head motions of the operator.

A concept of adaptive aiding for performance improvement in remote handling was described by A. Freedy, F. Hull and J. Lyman. A computer based system observes the operator's control function in relation to the environment and manipulator output, gradually takes over the manipulation task on a learning model based on maximum likelihood principles. This reduces the decision load of the operator to that of an action initiator and inhibitor. H.G. Stassen, A.W.A. Meyer and A. van Lunteren proposed a tactile display to transmit information between a prothesis and the wearer. Tracking experiments were performed to determine an optimal information exchange for the tactile display and a comparison with visual or auditory displays.

T.B. Sheridan described several current studies which have continued the investigation of supervisory control of manipulators, now termed tele-operators. One study, now in the hardware development phase permits a medical investigation of a remotely located subject by manipulation of a TV monitor-six-degree-of-freedom arm slaved to a similar arm-TV camera at the remote site. The second study is a more basic investigation of the man-computer interface where the computer assists the operator in the structural organization of sub-tasks to achieve an optimal solution of the primary tele-operator task.

V. ADVANCED MODELING TECHNIQUES

The emphasis of the papers presented in this session remained on the modeling of the human operator while tracking. However, two papers on supervisory control and a rather unique one on "urban dynamics" were presented.

R. Mekel and P. Peruo devised a model reference control system which permits the development of a nonlinear and adaptive model of the human operator. Extending the single axis model of Fu and Gould to two axis, D. Gilstad and K. Fu proposed a pattern recognizing, mode-switching model of the human operator in a visual-manual compensatory tracking task. J. Ware proposed an input adaptive model which weights error and error rate in the prediction of future error. Undaunted by last year's debate over optimal forms of the human operator, H.M. Paskin presented a discrete stochastic optimal control model of the human operator. The one variable parameter model has been tested for both pursuit and compensatory tracking over a range of input bandwidths.

Based on the results of a four display, four axis, two controller experiment, W.H. Levison presented a model for interference among multiple continuous manual control tasks. The model is able to predict performance scores and accounts for the effects of multivariable tracking on the controller's describing function and remnant.

R.W. Pew described two pilot experiments which will investigate two views of modeling the "system monitor and parameter adjuster". Theory suggests the task of parameter adjustment can be accomplished without the requirement for the motor and neuromuscular components. Practically, more effective parameter identification may result from a suitable man-computer interactive system than from either acting alone.

Also modeling the supervisor situation, T.B. Sheridan presented a procedure for specifying how long a supervisor or monitor of a process should wait between input samples to maximize a given value or payoff function. The procedure is based upon Bayesian preposterior information analysis.

Addressing himself to a seemingly vastly different modeling situation, E.S. Krendel presented criteria for an "adaptive urban system" based on a closed loop measurement of urban administrative response for a series of city departments. The study demonstrated the potential value of applying systems engineering techniques to the "quality of urban life".

VI. DESCRIBING FUNCTION MODELS

J. Heifferon and R.A. Hannen analytically verified the experimental results of Reid and Gordon-Smith which showed differences in operator performance for different spectral shapes of the input forcing functions.

A new adjustment rule for the analytical-verbal describing function model is suggested for continuous input spectra. During studies of a cyclist stabilizing a bicycle simulator, van Lunteren and Stassen found that the behavior of the rider is time-independent over at least five minutes. Also, the variation of parameters for one subject within one test, for one subject over a number of tests, and for a group of subjects has been determined.

G.A. Jackson investigated the optimal properties of the human operator when considering the crossover model in which the parameters were adjusted to minimize a cost function. The conclusion is drawn that all subjects optimize the same cost functional when given the same control task. The results of a current investigation of controller remnant were presented by W.H. Levison. Experiments have been conducted to explore the effects of display signal amplitude and bandwidth on the equivalent observation noise spectrum.

Research on a new human response test battery was reported by H.R. Jex and R.W. Allen. A battery of critical tracking tasks and step reaction-time tests was developed to permit efficient measurement of the limiting human dynamic response properties. Simple correlations are shown between the critical instability and various other closed-loop dynamic performance metrics. A number of psychophysiological measurements were taken while performing the test battery described above. These data showed consistent increases in the neuromuscular tension indicators during tracking. R.C. McLane described preliminary experiments seeking psychophysiological correlates of perceptual workload by aircraft pilots in performing a simulated two-axis flight tracking task.

VII. DESCRIBING FUNCTION MEASUREMENT TECHNIQUES

Three papers addressed themselves to the development or comparison of techniques for operators in manual control systems.

A paper by R.W. Allen and H.R. Jex presented an on-line Fourier analysis technique and two practical methods for mechanizing the technique. An electromechanical method was employed for the experiments described previously.

R.S. Shirley investigated the techniques of frequency analysis, parameter tracking, filtering and cross correlation for accuracy in measuring human operator frequency response. All methods provided good measurements in the region of crossover with varying accuracies away from crossover. All methods deteriorated when signal levels are low.

The last paper presented at the conference proposed a solution for the problem of excessive data scatter at low frequencies when using frequency domain techniques for pilot model measurement. L.W. Taylor suggested time domain analysis does not exhibit this scatter and further does not exhibit the low frequency phase lag indicated by the frequency domain analysis.

BULL-SESSIONS

The evening of the second day of the conference, the conference participants gathered in small groups for informal discussion of four "bull-session" topics. A brief summary of each of the group discussions follows:

Display and Manipulator Systems

The discussion opened with attempts to define just why the V/STOL display problem was different than other display problems. It was pointed out that since instrumentation is not usually available to provide a precise horizontal position reference that control of hover position must be based on rate information or on careful vehicle attitude control to prevent buildup of rate. The use of a cross-pointer display for ground position causes problems because the direction of motion stereotype based on ILS cross-pointers is reversed for the fore-aft dimension.

This led to a rather virorous discussion of the time to transition from looking out the window to reading instruments. It was asserted that when the problem was detecting a change in an instrument while monitoring the far field, detection times as long as 1.8 seconds were typical, but it was argued that this time was not applicable to the case of intentional transition to obtain a particular instrument reading.

The discussion then shifted to the question of manipulator design and ranged widely over topics including: (1) Questions of optimum control feel and the appropriateness of the matched manipulator concept; (2) The relative importance of proprioceptive feedback vs visual feedback, and laboratory studies directed at assessing their relative importance; and (3) The problem of how to design a four-degree-of-freedom control stick that retains some compatibility with vehicle degrees-of-freedom. This latter question was considered in the context of the V/STOL problem. Thus the group provided some closure with respect to the discussion topic if not with respect to the substantive issues under consideration. (Summary by Richard W. Pew, University of Michigan)

Performance Measures

The attendees were mainly interested in environmental effects on human performance, and as a result this proved to be the central theme.

The problems posed were: (a) how does one select a realistic environmental simulation (including physical environment and motivation), and (b) what should be measured to indicate performance. The former is an extremely complicated problem, with no general answer. Opinion ranged from a "partial" approach wherein everything else is held constant and some "stress" is increased until the subject "fails", to a plea for more realistic multi-stress experiments. No real conclusions were reached.

The second problem is somewhat simpler than the first in certain cases. Perhaps pilot "workload" and opinion rating can be found by measuring closed-loop performance and pilot model parameters. Or, the amount of remnant that results from more than one task may be a direct measure of workload. Also sub-critical task work in this area looks promising. In all, for simple tracking tasks several useful approaches to performance measure selection exist. However, it was also brought out that for many other real world manual control tasks, performance measure selection is not at all simple. Again, no real conclusions were expected nor reached.

Overall, a useful exchange of ideas, if not answers, took place. (Summary by R.O. Anderson, AF Flight Dynamics Laboratory)

Optimal Control Applied to Manual Control

The majority of discussion centered on the applicability of cost functions to manual control. Specifically, arguments were made about the use of absolute value of error as opposed to mean-squared error in the cost function. No concrete conclusions were reached on this subject. The question was also considered as to how optimal control can be applied to determine the controlled element dynamics such that when these dynamics are implemented with the human operator in the loop, the desired performance is obtained. Again, no conclusions were reached. (Summary by H.M. Paskin, AF Flight Dynamics Laboratory)

New Applications and the Future

The bull-session focused on new applications centered on large scale societal systems and criteria for defining (and in someone's utopian view) optimizing them.

It was agreed that the control loops of such systems tend to be interlaced, and the point of view (perhaps the source of inputs and observer of outputs at a particular location) determines the appropriate model. Expressed in another way, one man's inner loop is another man's outer loop.

The nature of transient response and what one can determine at early stages was another topic of discussion. By when can one tell if the curve is a positively increasing exponential as compared to a damped or saturating response? And against how much noise?

Much time was spent on the problem of who decides on criteria: when more than one person decides, how do you weigh their votes, and who serves as expert judge for what clients?

We all agreed that data are nice, and that data-based identifications are better than the other kind.

The session concluded without resolution on specifics of solution of the world's problems, but everyone seemed happy, and that's what counts. (Summary by T.B. Sheridan, MIT)

The Use of Predictor Displays for Terminal Area Air Traffic Control

by

William B. Rouse

Man-Machine Systems Laboratory
Department of Mechanical Engineering
Massachusetts Institute of Technology

Introduction

It is common knowledge that the Air Traffic Control (ATC) system is having problems, but the specific details of the problems and their sources are poorly understood. With this in mind, the Man-Machine Systems Laboratory at M.I.T. undertook a survey of the state of ATC[1] which included visiting airports and the FAA Experimental Center at Atlantic City (NAFEC) to talk with controllers, as well as an extensive literature search.

As might be expected, the results of this survey showed that ATC has several problems of various types and sources. The problems extend from those associated purely with engineering to financial and political considerations.

A Problem

One problem of concern to Man-Machine Systems is that of determining the proper role for the controller. Some solutions now being proposed include automation of the ATC system to the point that the controller becomes a passive and parallel element in the system. Proponents of such a solution, however, are quick to add that a controller is needed to run the system when unusual circumstances occur. Such unusual occurrences might include damaged planes in the approach pattern, stalled aircraft (A/C) on the runway, and pilots new to an airport and unfamiliar with the control system.

It appears that the controller cannot be subjugated to a standby role in ATC: he could not be expected to respond quickly and efficiently to emergency situations if he is not an active part of the system.

The solution seems to be the combining of talents of controller and computer, but the question of what the computer should do and what the man should do remains to be answered.

The Use of Predictor Displays for Terminal Area Air Traffic Control
William B. Rouse

The Study

Before presenting a plan for considering this man/computer question, it is important to be aware of the controller's present role.

Figure 1 is a sketch of a terminal area (TMA). An A/C may enter the TMA through one of several entry fixes which are defined by radio beacons. At these points, the controller of the regional control center "hands-off" the A/C to the TMA approach controller. The approach controller is aware that the A/C is due to arrive because he receives the flight plan of that A/C from its point of departure. Upon entering the TMA, the A/C can be either instructed to land or, if other A/C are waiting to land, to proceed to one of the holding stacks.

The controller calls A/C from the stacks on a first come-first serve basis. He instructs them to proceed to the regulated "funnel" and on to the gate or the beginning of the glide path. He attempts to schedule his calls to the stacks in such a way that a steady stream of A/C with separation intervals of three miles arrives at the gate. If the spacing is over three miles, he is wasting airspace. If it is under three miles, he is increasing the probability of a collision. The basic situation is that of an air traffic controller using his intuition and skill in giving speed and bearing commands to a number of A/C that approach from various directions, such that all of the A/C will merge into a single efficient stream of traffic.

To study a possible man/computer interaction for the air traffic control system, a predictor display was chosen as an aid for the controller. An ATC simulation (Figure 2) has been constructed to evaluate such a system. One analog computer generates A/C which are all "flown" via a control panel by a single operator. An analog/hybrid computer receives the "radar" signals and displays them to the subject. The subject's task is to perform the controller's role as previously discussed. To accomplish his task, the subject can use any of three modes of operation of the system. He can use a conventional system, one that predicts the paths of all A/C continuously, or a mixture of both such that some A/C appear conventionally and some appear with predicted paths.

The performance of the subject is measured using an index that penalizes him whenever the spacing of A/C is other than three miles. His performance will be compared to data taken using solely the conventional mode of the system. The subject will be rewarded according to his performance.

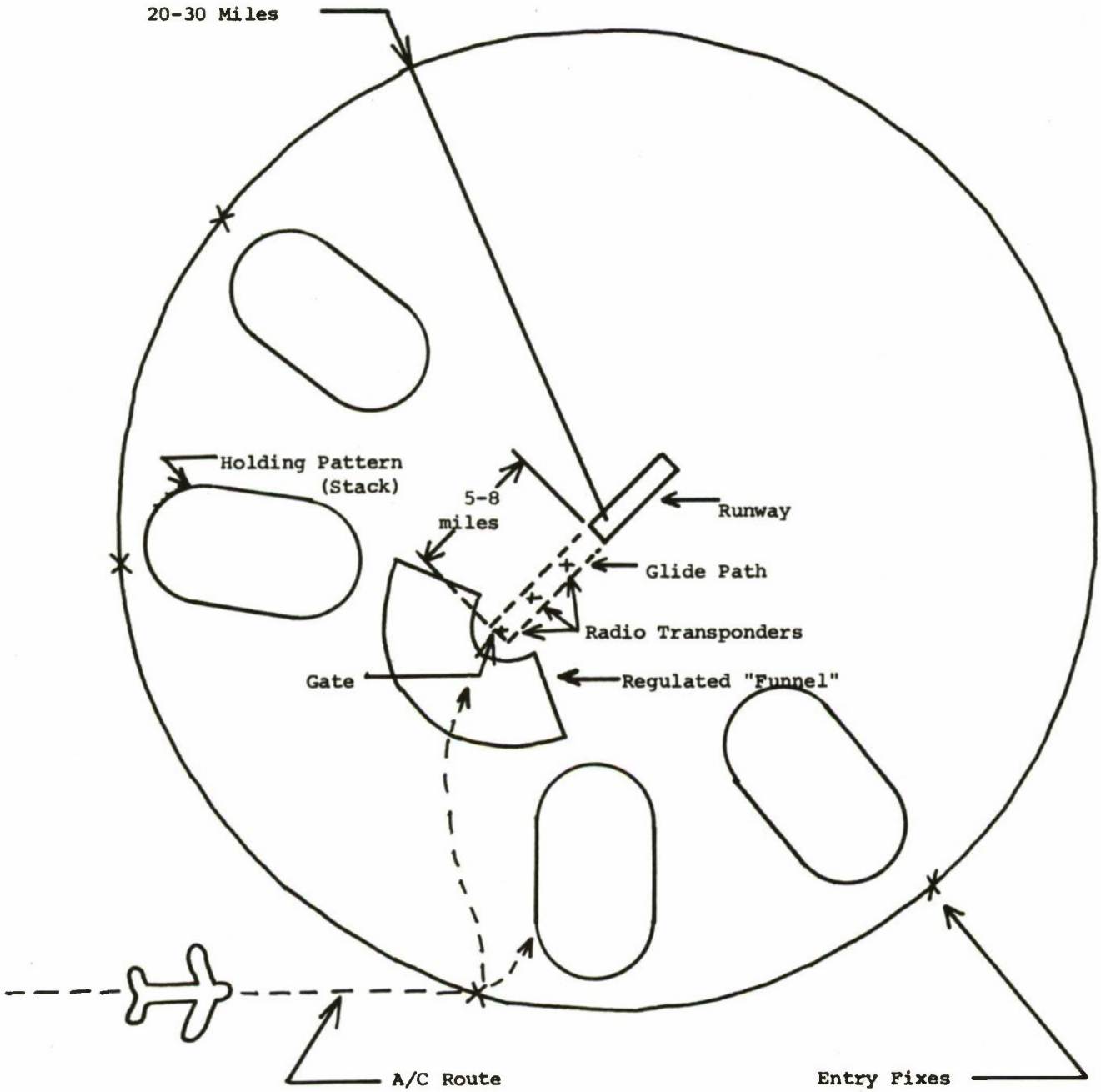
The above system is now operational, but as yet no data has been accumulated.

Summary

The purpose of this research is to investigate the use of a computer to aid a subject in performing a complex task. The ability of the computer to rapidly make calculations in updating predictions will be coupled with the flexibility of the human response. The conclusion of this work should give an indication of how the above man/computer system for air traffic control would compare with the conventional system. Further studies may show how such a system would compare with one that allocates more responsibility to the computer.

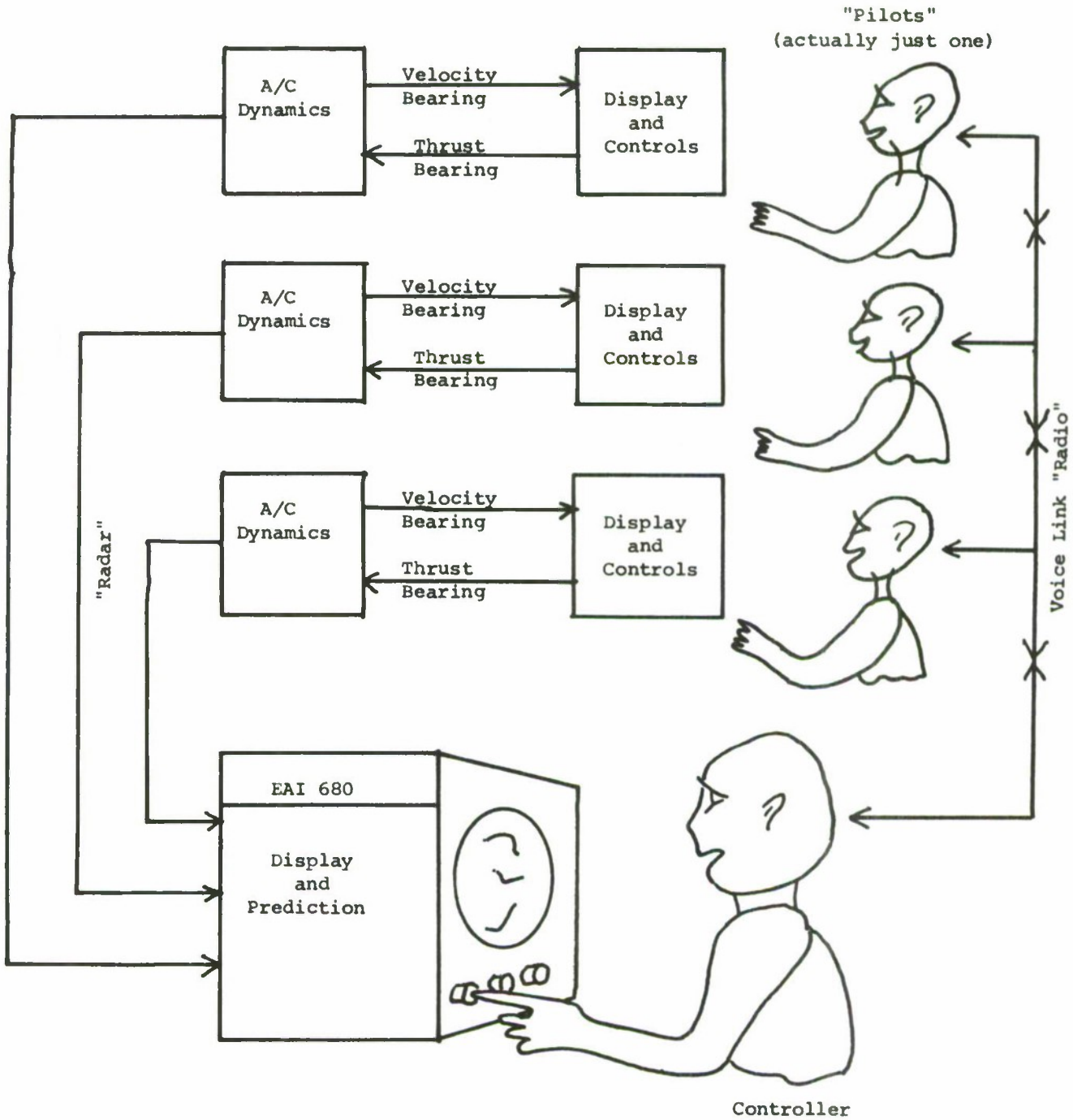
Reference

Rouse, W.B. An Appraisal of Problems in the Air Traffic Control System, Engineering Projects Laboratory Report No. 70283-12, M.I.T., March 1970.



Terminal Control Area

Figure 1.



Air Traffic Control Simulation

Figure 2.

AN INVESTIGATION OF HEAD TRACKING PERFORMANCE USING A HELMET MOUNTED SIGHT AND DISPLAY SYSTEM

Allen H. Henke
Honeywell Systems and Research Center

An investigation of head tracking behavior was accomplished using a helmet mounted sight and display. The purpose of the study was to gather basic head tracking data and to demonstrate the feasibility of a helmet mounted display.

The structure of the experimentation was evolved by considering the application of a helmet mounted device to the pilot task of a single place high performance fighter aircraft. The study was funded by the Flight Dynamics Laboratory at Wright-Patterson AFB under the direction of Mr. Eldon Bobbet. (The simulation was not specifically designed to portray the complete pilot task, but was designed to provide basic data for further hardware development and simulation).

The principle advantage of a helmet mounted sight and display system is that it is possible to provide a pilot with both a unique and versatile source of head-up information along with a new dimension of control freedom. Simply providing the pilot with more display information without providing him with a new control channel can only increase an already bad situation.

As we all know, the early developmental stages of most new system concepts are plagued with the same affliction; i.e., very little data exists to document how well the system can be used by the operator. Intuitive reasoning may suggest the concept is correct and a trained

operator could use it, but quantitative and inferential statements cannot be made until data have been gathered.

For the most part, the available tracking literature can only be of use in the intuitive portions of this system analysis. It can be shown that the human operator can track a moving target using eye/hand coordination and it can be accomplished quite accurately for simple non-aided systems. However, the strategy which an individual employs in a tracking task may be significantly different and result in performance changes if he is required to track with a different response mode. For this reason, basic head tracking data was collected to determine how accurately an operator tracks moving targets using the large and small muscles of the head and shoulders, and whether these data are comparable to present tracking accuracy descriptions.

It is reasonable to assume that human operators are experienced in tracking moving targets with their heads and eyes. The complex and rapidly moving systems in our environment force an individual to become proficient in estimating closure rates and angular movements of assorted targets. Furthermore, an individual's own movement with respect to his environment creates a head and eye tracking situation if he is to avoid other objects and move about freely. The ability of an operator to track a moving target using eye/head coordination is basic to a number of tasks. Driving a car, making a final approach to landing or scoring a "hit" in air-to-air combat, all involve head tracking behavior to align his visual system to a target or object.

Although a great deal of emphasis has been placed directly on increasing the accuracy of this behavior, direct performance measurement has been of little operational interest except as an integral part of other training protocols. However, recent advances in display technology and methods for measuring head movement have lead to a need for documenting response accuracies associated with head/eye coordinated tracking performance.

The present experiment was designed to produce preliminary information concerning head tracking accuracy and display usability. Questions concerning tracking accuracy were of primary interest along with the measurement of target acquisition times and the effect of providing cueing information to the operator. The concept of presenting remotely displayed information via helmet mounted display was implicit in the apparatus design.

Experimental Design

A 3 x 2 mixed analysis of variance design was used to combine three levels of target dynamics and two types of reticles. The target types were: 1) static; 2) predictable path; and, 3) evasive. The static targets appeared within the target area and remained stationary. Predictable path targets entered the target area and moved in either a straight line from one quadrant into the opposite quadrant or in a slightly curved line from one quadrant to the opposite quadrant. The evasive target paths were defined as those which made three or more changes in direction. In all cases, the intitial display position of the targets was randomized with respect to the quadrants of the target

area. The target rate was held constant at 6° per second for both the predictable path and evasive target conditions. Two reticle types were chosen to provide design information into the relationship between orthogonal and triangular reference marks and tracking accuracy.

The subject's tracking performance was evaluated as a function of the root mean square of radial error (RMS_r), mean radial error ($\bar{R}$) and time measures. The following time measures were recorded for each subject:

Reaction time - time measured from onset of cueing signal until the subject moved his head 2% of the total target area.

Target acquisition time (TAT) (actual) - time measured from the subject's initial head movement until the reticle was placed within a 1-1/2 x 1-1/2° tracking gate around the target.

Target acquisition time (TAT) (indicated) - time measured from subject's initial head movement until the target acquisition switch was pushed to indicate that the subject estimated he had acquired the target to his best level of accuracy. (It was anticipated that indicated target acquisition time would prove to be longer than actual target acquisition time due to the tendancy for the subject to initially overshoot the target and return to it.)

Time on target (TOT) - time measured while the subject was within a 1-1/2 x 1-1/2° gate surrounding the target. Both the RMS_r and $\bar{R}$ errors were computed for the time prior to and after indicated target acquisition time.

A Sanborn recorder was employed to record the tracking profiles from the subjects under each experimental condition. Analog voltages of the helmet sight system were recorded for elevation and azimuth. A continuous recording of radial error was included to provide a display of momentary accuracies which would not be directly apparent in the summary data from the printer.

Each of the six subjects were required to perform under all six experimental conditions for a total of 36 cells. Each cell was composed of 40 trials for a total of 1,440 trials (240 under each condition). A trial was defined as a 15 second target presentation beginning with the appearance of a sighting cross on the display and terminating with the disappearance of the total display. The order of presentation was counterbalanced to minimize order effects across subjects. An attempt was made to counterbalance time of testing. However, due to other working commitments, the subjects had to be randomly assigned to the testing schedule. The data were gathered between the hours of 0800 and 1100.

The subjects were six male systems engineers employed at the Honeywell Systems and Research Center. The ages of the subjects ranged from 31 to 44 with a mean age of 38. The only criterion for subjects selection was that the subjects did not wear corrective lenses. Although they were aware of the helmet sight operation, they had not been exposed to the display system prior to entry into the experiment. The subjects were briefed on all aspects of the experiment including the equipment design prior to entry into the experiment.

A closed loop simulation was designed which provided for the remote display of computer addressed information on a helmet mounted one-inch cathode ray tube. The output of the helmet mounted sight system was fed through A to D converters to a digital computer. The computer was programmed to drive a 19 inch display scope containing a sighting reticle, cueing signal, and target. The helmet mounted CRT was made compatible with the output of the digital computer so that all information presented on the 19 inch monitor scope was also available on the helmet mounted display. By adjusting the helmet mounted CRT amplifications, any degree of magnification was possible; i.e., the helmet mounted CRT could display the entire 19 inch display or coverage could be reduced to present any percentage of the total scope area.

For purposes of the present study, only the center portion of the total area was displayed so that all targets initially appeared out of the subject's field of view. A cueing signal consisting of an arrow pointing at a 45° angle into the quadrant where the target initially appeared was used to direct the subject's head movement to the target.

The known position of the helmet in elevation and azimuth was combined with the programmed position of the target to measure and display the tracking error to the subject. Therefore, the target was displayed at a position which was representative of the error in azimuth and elevation between the output signals of the helmet and the programmed position of the target. If the target was tracked so that the position of the helmet equalled the position of the target, the error signal in azimuth and elevation was zero and the target appeared in the center of the reticle.

Procedure

Briefly, the procedure for each trial was comprised of the following steps:

1. The sighting cross appeared, and the subject centered the reticle over the cross.

2. The cueing arrow appeared pointing to the quadrant containing the target and the cross disappeared.

3. The subject quickly swung the reticle into the active quadrant and positioned the reticle pipper in the center of the target.

4. When the subject felt that he had acquired the target to the best of his ability, he pushed the target acquisition switch.

5. The target was then tracked for 15 seconds whereupon the trial terminated and the display imagery disappeared.

6. Following an intertrial interval of 15 seconds ($\pm$ 3 seconds) the sighting cross appeared, and the next trial began.

7. After completion of 20 trials, the subject was given a 5 minute rest period and was allowed to leave the testing cubicle.

The above steps were followed for the remainder of the testing session.

Results

A simple linear regression analysis of variance was performed between all pairs of appropriate dependent variables to insure that the expected statistical relations were supported. Table 1 presents the obtained correlation coefficients. A high positive correlation exists between the root mean square of radial error and mean radial error across 1440 trials. As expected, a high negative correlation exists between both RMS_r and $\bar{R}$ versus time on target.

A 3 x 2 mixed model analysis variance was performed on the average radial error data. For the two fixed effects, only differences between target types was obtained. ($P<0.01$) indicating that although the three target types differed in difficulty, the reticle variability did not significantly affect the accuracy of tracking behavior. A significant difference ($P<0.01$) was also obtained for the subject's randomized effect. The interaction between target types and reticles was not significant. Similar results were obtained on time on target data. The $\bar{R}$ and time on target data and standard deviations for each condition are presented in Table 2. The order of increasing difficulty was static, predictable path, and evasive.

In all cases the mean actual TAT measurement as detected by the computer was shorter than the indicated TAT by the subject's response.

No significant differences were found for the indicated TAT data for all conditions. Reaction time to the cueing signal was 0.519 seconds.

Discussion

The usability of a helmet mounted sight and display concept was clearly supported. An operator equipped with such a system is able to track both static and moving targets with a relatively high degree of accuracy. However, larger mean deviations from target center can be expected to occur as a function of target movement.

Increases in tracking error from static to evasive targets was assumed to be a function of the operator's inability to anticipate changes in target direction and to establish a steady rate of head movement.

The static target, once it had been acquired, was completely predictable. Errors associated with the target condition were simply related to system "jitter". As defined, the predictable path targets required the subject to establish a constant rate and direction of head movement. Thus, predictable path tracking errors reflect the subject's inability to maintain a steady rate of head movement in a specific direction. Evasive targets added yet another dimension to the operator's performance error. Since the target was programmed to make three or more changes in direction, the subject's were unable to predict the final target course and therefore were forced to chase the target whenever it changed direction.

Although the orthogonal and triangular reference marks did not affect tracking accuracy, the subjects quickly complained if the center dot or pipper was not clear or bright enough. It would appear that the subject's relied upon the center mark for precise positioning of the reticle over the target. Subject comments supported this assumption.

The consistently longer TAT indicated by the subjects was primarily due to tendency to "overshoot" the target during acquisition. An inspection of the continuous recording data revealed that in all cases the subjects were unable to stop on target center, but passed through the target and returned. The indicated TAT measurement reflected the time to move to the target plus the time to make minor centering adjustments, whereas the actual TAT measurement represented the time taken to move the reticle pipper from the center of the target area through the target center.

Source	r
RMS_R versus $\bar{R}$	0.99
RMS_R versus TOT	-0.91
$\bar{R}$ versus TOT	-0.95

Table 1. Correlation Coefficients Computed Between RMS_R Mean $\bar{R}$ (R), and TOT

Dep. Measure	Static Ret 1	Static Ret 2	Pred Path Ret 1	Pred Path Ret 2	Evasive Ret 1	Evasive Ret 2
Mean Radial Error	0.680	0.682	1.054	1.027	1.198	1.216
Standard Deviation	0.135	0.145	0.333	0.275	0.261	0.375
Mean Time-On-Target	11.700	11.695	6.953	6.820	5.410	5.387
Standard Deviations	2.540	2.712	2.230	2.316	1.980	1.828

Table 2: Mean Radial Error (in degrees) and Time-on-Target (in seconds) and Standard Deviations for all Test Conditions

CONTROL GAINS IN HEAD-UP PRESENTATION

J. M. Naish

Douglas Aircraft Company
Long Beach

Recent flight and simulator tests of a head-up display system for civil aviation are reported. An optimized format of conformable command and attitude symbols is used to track ILS beams in the latter part of an approach, during which human performance is assessed by objective and subjective techniques. Tracking accuracy is shown as a function of seven control gains for real and simulated flight in a DC-9-20 aircraft, and for a small number of subjects.

The results are used to optimize gains, by means of which a similar level of performance is achieved by a larger group of subjects. Human tracking is found to be as good as that of an automatic pilot, and the manual head-up approach is compared with automatic operation, making use of additional results from studies of learning and the transition from instrument to visual flight. It is shown that a manual approach with HUD can allow the pilot greater command of the relevant information than is possible in automatic flight.

LIST OF ILLUSTRATIONS

1 Display Format

2 Overhead Installation Showing Head Clearance and Unobstructed Windshield

3 Azimuth Command Computation

4 Elevation Command Computation

5 Lateral Error and Cooper Rating Against Heading Gain

6 Lateral Error and Cooper Rating Against Bank Rate Gain Ratio

7 Vertical Error and Cooper Rating Against Elevation Gain

8 Vertical Error and Cooper Rating Against Elevation Rate Gain Ratio

9 Lateral Error and Cooper Rating Against Heading to Bank Gain Ratio

10 Vertical Error and Cooper Rating Against Glide Slope Gain Ratio

11 Lateral Error and Cooper Rating Against Localizer Gain Ratio

12 Distribution of Lateral Errors for Visiting Pilots

LIST OF TABLES

I Method of Surveying Gains

II Optimum Gains for DC-9 Aircraft

III Comparison of Automatic and Manual HUD Tracking

IV Comparison of Manual, HUD and Automatic Approach Systems

CONTROL GAINS IN HEAD-UP PRESENTATION

This paper is part of a more comprehensive report on recent flight tests of the Head-Up Display (HUD). It is concerned with improving an aspect of head-up presentation which has not received adequate attention. Previous practice[1] has been based on the manipulation of gains to suit the opinions of a limited number of highly experienced subjects, in the hope (largely justified by events) that resulting values would be suitable for subsequent users. Such an ad hoc method was necessary when the main concern was with removing gross effects hindering the acquisition of visual information from display and forward view, effects mainly graphical in nature; but having removed these hindrances[1] it is desirable to arrive at a more systematic understanding of the time-dependent aspects of the system, by methods placing less reliance on subjective techniques.

It is still advisable to work with a small number of pilots, at least in the airborne part of the investigations, because the economic aspects of flight testing cannot lightly be dismissed. The consequent need to establish generality is then to be satisfied by applying the results to a larger population of subjects. It is also desirable to correlate flight test results with results obtained by simulation, with a view to estimating satisfactory gains from ground studies. The work to be described is thus in the nature of a gain survey, using a small number of subjects and an approach not limited to subjective methods of assessment, with the intention of relating to a larger group of subjects and simulation methods.

Adaptation of Display Dynamics to Fit Pilot

The display format contains symbols representing chosen items of information, which are drawn according to graphical and organizational rules designed to help the user understand their meaning[2]. The display symbols, especially the director symbol, also move or change shape according to dynamic rules, and these are governed by gains applied in control loops operating on the inputs from data sources. When the pilot acts on information conveyed by the changing symbols, he closes the control loop, and there is a similarity between a directed manual approach and an automatic approach, in that both may be considered as operations on signals originating in aircraft sensors. It follows that flight director gains might be adjusted as autopilot gains are adjusted, for best performance, but in this case seeking to adapt display dynamics to suit human rather than automatic capabilities.

The basis for optimizing gains could be a theoretical model of the control situation, in which the pilot is represented by describing functions. An investigation is currently being carried out by M. Abramovitz in which this combination of human operator and aircraft is analyzed, and gains are optimized by means of Bode plots and root loci, starting with the inner loops. In the present investigation, gains are examined on an entirely experimental basis. An optimized gain is expected to allow minimum error in tracking the director symbol, in the command channel affected by the gain, because the pilot should then be able to follow commands most continuously and accurately. On the

other hand, minimum tracking error may not occur at the gain giving the best subjective rating; for example, because the pilot is required to make unacceptable control movements. Gains are therefore optimized also with respect to user opinion about handling quality in the affected channel, and for this purpose the Cooper Scale provides a convenient basis of measurement. While optimizing a particular gain, other gains are held at values suggested by preliminary studies for the type of aircraft used in the flight tests, and these gains are shown in the first two columns of Table I.

EXPERIMENTAL ARRANGEMENT

The measurements to be reported were obtained in a flight test program with a DC-9-20 aircraft, and in a similar program carried out subsequently with a simulation of the same aircraft. To this end, the entire experimental assembly, for generating and displaying command information and also for measuring tracking accuracy, was designed to be transferred easily between laboratory and flight deck.

Airborne Equipment

The display format and installation, optimized in other studies[1, 3], are shown in Figures 1 and 2, respectively, from which the pilot's visual field and immediate environment may be understood. During experimental runs, the external view was blanked off by covering the main forward-facing panel of the windshield with a sheet of polaroid, and by subjects wearing (crossed) polaroid goggles, through which only the display could be seen. Mode annunciator and display control facilities, though playing no direct part in the recorded sections of experimental runs, were placed close to the reflector plate.

The flight director commands were generated in an experimental analog computer supplied by Sperry, Phoenix, for which the block diagrams are shown in Figures 3 and 4, for heading and elevation channels, respectively. The diagrams have been drawn without some devices used for limiting, washout and filtering, thus considering only the dynamics within the frequency range relevant to the tracking task; namely, between about 0.2 and 6 radians per second. It will be seen that, in addition to the usual mixture of a glide path signal and an attitude signal, an attitude rate term has been included for the purpose of providing a more immediate display response on changing attitude.

The method of expressing gains was based on relating angular displacements of the flight director index, γ_A, γ_E, to the Euler angles of attitude, ϕ, θ, ψ, and to angular deviations from the glide path, σ_A, σ_E. For example, heading gain, K_ψ, was derived from the azimuthal deflection of the director index in the collimated display field and the associated change of heading input. In computing gains, the total input signal was used; that is, contributions were included from paths omitted in Figures 3 and 4, so that computed gains did not coincide with the static gains (for example, in cases where washout circuits eliminated the very low frequencies).

Tracking errors were integrated between heights of about 1200 feet and about 250 feet, during which period the flight director computer was used in a fixed mode with selected, constant gains. The error integrating unit was designed

to yield the mean absolute, or mean modulus, error in azimuth and elevation during experimental runs of about 100 seconds. Its position in relation to the director computer and Head-Up Display circuits is shown in Figures 3 and 4. Error scores in the two command channels were calibrated in terms of equivalent angular glide path errors, $\langle \sigma_A \rangle$, $\langle \sigma_E \rangle$, having fixed values and giving the same error score during the same interval of time.

Simulator Equipment

At the conclusion of the flight test program, the experimental equipment was transferred to the simulator and used with very little modification for the laboratory program. The collimator was now clamped directly onto the pilot's side of the instrument panel and the display format was presented against a dark background by a temporized reflector plate, mounted at 45 deg to the forward sight line. Mode annunciator and control facilities were provided in the vicinity of the reflector, without any requirement for precisely copying the airborne arrangement. Conventional flight instruments were not used for the experimental task but were included in the experimental rig to maintain similar electric loading on data and computer sources. The flight director computer and the error integrator were used without change, but the measuring equipment was re-calibrated at this stage. Standard analog methods were used to simulate a DC-9-20 approaching at 1.3 Vs + 5 kts , with slats extended, gear down, and flaps deflected 50 deg.

EXPERIMENTAL METHOD

Airborne Test Program

The experimental runs were made mostly by one subject, S1, because the total number of runs was limited, but in one case ($K\sigma_E / K\theta$) another subject (S2) was used. The basis for experimental procedure is shown in Table I, where each column includes a set of gains held fixed while varying the gain chosen for investigation, V. It was not necessary to take the variable gains in strict sequence, from the inner loops ($K\dot{\phi}$, $K\dot{\theta}$) outwards, because most gains were already set at nearly optimum values. However, some values were changed to suit subjective rating: notably, rate gain ratios were reduced nearly fourfold and elevation gain, K_θ , was reduced by a factor of two, the latter having some effect on error score. For each gain investigated, it was usually possible to choose seven values, each being flown for one approach. In this way, two or three gains were surveyed during each test flight of three to four hours duration.

The experimental task was to fly an approach, between the prescribed heights, while correcting the small path errors arising through beam noise. The approaches were flown at Ontario, Bakersfield, March Air Force Base, and Stockton, California, mostly in smooth air and with autothrottles engaged. There were no practice runs, or replicated runs, because the learning effect was known to be negligible for the format of Figure 1[4]. Values of the variable gain were taken in random sequence, and these values were unknown to the user.

Simulator Test Program

The same subject, S1, made the experimental runs in the laboratory test program. Gains were investigated in the same order as in the flight test program, Table I, with the same sequence of (unseen) values for the variable gain, V. Two or three gains were investigated in each experimental session, which was only of about an hour's duration because no time was needed to reach the experimental venue or to fly downwind legs. Some practice runs were flown to make up for this.

The experimental task was again to correct path errors due to beam noise during an approach in smooth air between the same prescribed heights. The beam noise was set to give the same subjective impression as in real flight conditions. For this purpose, it was sufficient to inject at the glide slope receiver a noisy signal having an r.m.s. amplitude of 30 mv, corresponding to an r.m.s. path error of 0.14 deg, and flat to 30 cycles/second. The corresponding figures for the independent noise input to the localizer receiver were 3.8 mv, or 0.056 deg path error, and 1.5 cycles/second. All runs were performed without the help of autothrottles and the subject was thus loaded with an auxiliary task, not called for in the flight tests.

After concluding the experimental investigation of control gains, values were selected for use in two series of demonstration flights, for the purpose of establishing a connection with a larger group of subjects. The first series was for company pilots, with display gains set to the values shown for S3-5 in the lower part of Table II. These are the values reached during the gain survey, which are also shown in the last two columns of Table I (except that a range of values was used in programming the glide slope gain ratio, $K_{\sigma E}/K_{\theta}$, below the limiting height of 250 ft.). They were mostly lower than the experimentally determined values shown in the preceding rows of Table II, but it can be seen from the data plots, which are to be presented in Figures 5 to 11, that the corresponding change in error score would only exceed 0.01 deg for the changes in heading and pitch gains, and then only in simulated flight. The reductions, which were intended to give smaller symbol movements (and were required by other pilots using the system), were therefore made at the expense of a small performance penalty. The second series of demonstration flights was for visiting pilots, with gains set to the values shown for S6-41 in Table II. These gains had been yet further reduced, by 10 percent for heading and elevation and by about 25 percent for localizer ratio, because most of the visitors were without prior experience of the aircraft and were expected to feel more at ease with smaller symbol movements.*

The demonstrations consisted of three pilot-controlled approaches to a height of one hundred feet, in the same DC-9-20 aircraft. The first approach was generally flown with less than five minutes familiarization, acquired on the downward leg, and with forward visibility provided. The second and third approaches were made with the forward visibility reduced by the crossed polars to any value between a few hundred feet and one mile, depending on external conditions. On the last approach, a touch and go landing was to be attempted if the aircraft was suitably placed.

Tracking errors were again recorded, but in computing the equivalent angular offsets from the beam, in azimuth and elevation, the glide slope gain was assumed to be fixed (at an equivalent value). The measured runs were between

* The gain program was started at the outer marker in these runs.

approximately 1200 feet and 250 feet, as before. Approaches were flown by day and by night at March Air Force Base, Oakland, Fresno, Palmdale, Edwards Air Force Base, and Long Beach, California, in smooth air and usually with autothrottles engaged.

One of the approaches under reduced visibility was used to investigate disorientation. For this purpose, the display format was slewed in azimuth, so that the pilot would be confronted with a potentially disorientating situation when the external world became visible.

RESULTS

Results of the experimental runs are shown in Figures 5 to 11, inclusive, as plots of tracking error and Cooper rating against the gain or gain ratio under investigation. The results always relate to the command channel affected by the gain change; for example, heading gain, K_ψ , is plotted against lateral error, $\langle \sigma_A \rangle$, and against Cooper rating for the azimuth channel, in Figure 5. Error scores for the other channel are not shown because they were always found to be invariant within experimental limits.

Curves are drawn on the assumption that human performance and subjective evaluation, under the experimental conditions, are single-valued and continuous functions of the gain varied, except for chance effects. By inspection, values are then selected which reflect the best operating conditions according to each method of assessment, with the results shown in Table II for real and simulated flight. In some cases, lowest scores are approached asymptotically and the gains shown are then values at which performance starts to deteriorate.

Heading Gain, K_ψ , Figure 5. Variation of heading gain in simulated flight caused tracking errors to increase for values of the gain less than 0.15, and the best Cooper rating was given at about 0.125. The gain was varied over a smaller range in real flight, yielding less information. There was no recognizable trend in error score, and the mean error, of 0.056 deg, was larger than the asymptotic minimum of 0.033 deg for simulated flight. Cooper ratings decreased to a possibly stationary value of 3 in real flight, which was the same as the lowest rating given in the simulator and might therefore be a true minimum, occurring at a gain of about 0.075.

Bank Rate Gain Ratio, $K_{\dot\phi}/K_\phi$, Figure 6. Variation of the bank rate gain ratio caused distinct error minima in both experimental situations. Lateral error was least at a value of 1.25 in real flight and at 1.5 in simulated flight, where the general level of tracking error was somewhat higher. Cooper ratings were least at a gain ratio of 1.25 in simulated flight. In real flight, subjective evaluation was best for ratios less than 1.0.

Pitch Gain, K_θ , Figure 7. Each measure behaved similarly in real and simulated flight. Tracking errors started to increase at pitch gains less than about 0.3, from a plateau having a height of 0.027 deg in real, and 0.018 deg in simulated flight. Cooper ratings were least for a pitch gain of 0.2, in both cases.

Pitch Rate Gain Ratio, $K\dot{\theta}/K_\theta$, Figure 8. The measurements taken in simulated flight showed a minimum tracking error occurring at a pitch rate gain ratio of 1.0, at an error level of 0.03 deg. In real flight, minimum error scores on the order of 0.04 deg were obtained at gain ratios less than 1.0.

Subjective evaluation in simulated flight showed a possible minimum at a ratio of about 0.8 which was less discernible than in real flight, where the preferred ratio was about 1.4.

Heading to Bank Gain Ratio, $K\psi/K\phi$, Figure 9. In-flight measurements taken while varying heading-to-bank ratio showed no trend away from a uniform performance level of 0.04 deg. Simulator results showed a slight upward trend in lateral error score, from the 0.02 deg level, at gain ratios greater than 1.2. Similarly, Cooper ratings were uniform throughout the experimental range of heading-to-bank ratio in flight, whereas in simulated flight there was a shallow minimum at about 1.2. The airborne results were obtained in slightly choppy conditions.

Glide Slope Gain Ratio, $K\sigma_E/K\theta$, Figure 10. Vertical error in flight, for subject S2, decreased slowly throughout the experimental range, reaching a steady value of about 0.03 deg at a glide slope gain ratio of 17. In simulated flight, a similar trend was found, for subject S1, with vertical error again becoming steady at a gain ratio of about 17, at a level of 0.02 deg. On the other hand, Cooper ratings were found to increase slowly as errors decreased, both in simulated and real flight situations. In each case, ratings levelled out at gain ratios less than about 15.

Localizer Gain Ratio, $K\sigma_A/K\psi$, Figure 11. Variation of the localizer gain ratio caused similar effects in real and simulated flight. In each case, lateral error decreased to a steady value at gain ratios exceeding about 40. The level of the plateau was 0.045 deg in real flight and about 0.025 deg in simulated flight. Cooper ratings were uniform throughout the gain range and equal in both experimental situations.

Tracking Accuracy in Demonstration Flights. Results obtained in the demonstration flights with succeeding groups of subjects are conveniently expressed by the lateral tracking errors, which were found greater than vertical errors. Lateral errors are shown as histograms for each of the three approaches in Figure 12, where smooth air results for 3 company pilots, S3-5, are shown hatched, and results for 36 visiting pilots, S6-41, are shown in full shading. The lowest scores achieved by S1 in the gain survey are represented by the vertical dotted lines, at the values 0.03 deg and 0.055 deg. These values are the levels of the plateaus to which error scores descended, the lower level occurring while investigating bank rate gain ratio, Figure 6, and the upper while investigating heading gain, Figure 5, in real flight and in smooth air.

Inspection shows that error scores for S3-5 were either within or below the limits for S1, on all three approaches. Of the visiting pilots, 69 percent returned error scores within the S1 limits on their first run, 70 percent had

scores within these limits on their second run, and on their third run 61 percent of the scores fell within S1 limits. The best scores of an experimental test pilot who was well acquainted with the system could thus be approached by a majority of visiting pilots using similar, but somewhat downgraded, gains. The same scoring level could be reached by all subjects in the company test pilot group, using optimized gains.

The results for the demonstration flights are also shown in relation to Category II limits by the chain-dotted line in Figure 12, which has been drawn at an error value of 0.13 deg. This value was calculated from the standard deviation of the tolerance permitted in lateral error between heights of 300 ft and 100 ft[5], as a mean permissible error. It is seen that all lateral errors, except one during the first approach, lie within the Category II limit; that is, the failure rate was less than 1 percent. The supervising pilots S1 and S2 also reported independently that all approaches were judged to be within Category II limits. Moreover, touch-and-go landings were made on the third approach after restoring forward visibility at a height of 100 feet. Deliberate misalignment of the format was found to be without effect on orientation, or on runway acquisition.

DISCUSSION

Optimized Operating Conditions

Considering first the experimental runs obtained in simulated flight, it is seen that most results lie close to a smooth curve drawn through the data points. Exceptionally, there is one wild point in lateral error plot of Figure 6, that is, a point removed by more than an experimental error of about 0.005 deg, and there are two in the lateral error plot of Figure 9. These exceptions are sufficiently rare to allow the inference that tracking performance could be regarded as a continuous, single valued function of a control gain. This view was reinforced by the fact that gain values were unknown to subjects, who also found it impossible to distinguish empirically between one control gain and another. The results obtained in real flight were somewhat less consistent, with departures from the error curve of nearly 0.01 deg in Figures 5 and 6. This difference is attributed to the greater difficulty of preserving uniform operating conditions in real flight. As regards Cooper Ratings, it will be seen that experimental results fall mostly within hàlf a point of a smooth curve drawn through the data points, and this amount of variation appears to be reasonable in a subjective measure. Both measures of performance were evidently functions of the control gain and this view is supported by the observation that scores were invariant in the command channel bearing no theoretical relation to the gain being varied.

The effect of adding another task in simulated flight, where subjects were required to operate the throttles, was evidently without adverse effect on performance of the tracking task. In all cases but one the level of tracking error was lower in simulated flight than in real flight: the exception is seen in Figure 6, where lateral errors were greater in simulated flight. For this reason the simulated flight results probably reflect relationships between performance and gain better than the real flight results, where there were evidently influences, such as turbulence, tending to mask the effects under investigation.

It was unfortunate that the same subject could not be used in all experimental runs and some degree of consistency has been lost for this reason. The effect of changing subjects was evidently very small, however, as can be seen from the results shown in Figure 10, for variation of the glide slope gain ratio. After allowing for increased error in flight, comparison of the performance levels and trends in the two experimental situations indicates that no great confounding effect was introduced by the change of subject, a view which was supported by subsequent comparisons of performance for S1 and S2.

The inference that the experimental results revealed relationships between performance and gain is further illustrated by a comparison of results obtained in simulated and real flight. On this basis, the results collected in Table II show the same value for the control gain giving minimum tracking error in four out of seven cases, viz, elevation, elevation rate, glide slope and localizer (columns 3, 4, 6, 7). In two of the other cases there was no variation in tracking performance in real flight, columns 1, 5, and in the case of bank rate (column 2) the difference for the two kinds of flight, of 0.25, was smaller than would correspond with an effect of chance. From this comparison it was concluded that the same relationship between gain and performance was shown in real and simulated flight, except in cases where the range of gain values was insufficient or where masking effects, possibly due to turbulence, concealed trends in performance.

TABLE I. METHOD OF SURVEYING GAINS

V	0.088	0.088	0.088	0.088	0.088	0.088
1.25	V	0.67	0.67	0.33	0.33	0.33
0.284	0.284	V	0.141	0.141	0.141	0.141
1.25	1.25	1.25	V	0.33	0.33	0.33
1.5	1.5	1.5	1.5	V	1.3	1.3
15.2	15.2	15.2	15.2	15.2	V	15.2
33.6	33.6	33.6	33.6	33.6	33.6	V

V signifies variable gain

Comparison of the optimum gains found by the two methods of measuring performance shows two kinds of effects, which are shown in the experimental curves better than in the tabulated results. In most cases the tracking error curve is seen to be roughly parallel with the Cooper rating curve, with minima at nearly the same gain or gain ratio. In the case of heading gain, Figure 5, and elevation gain, Figure 7, an entirely different effect is found: the curves have different shapes, and their minima are not coincident. A similar but less marked effect is seen in the curves for glide slope and localizer gain ratios, Figures 10 and 11. It follows that the objective method of measuring performance cannot always be used to give the best operating conditions, but should serve rather as a basis for cooperative adjustment.

The results obtained in varying elevation gain, Figure 7, best illustrate how the objective and subjective methods of measurement can be used to understand the control process and arrive at best operating conditions. The objective method shows that elevation gain cannot be reduced much below a value of 0.3 without increasing tracking error, in both kinds of flight. But the user prefers a value of 0.2, at which tracking error is about 50 percent greater, in simulated flight. The reason for preferring the smaller gain appears to be that the pilot then has the feeling of a less busy control situation, with less tendency to overcontrol in turbulence, and it can be seen that in almost all cases the gain preferred by the user was less than the minimum revealed by tracking performance. In choosing values likely to give the most satisfactory operating conditions it is necessary to strike a balance between the requirements for satisfactory handling qualities and small tracking errors. With this in mind it can be understood why gains were generally reduced for the demonstration flights, at the expense of a very small performance penalty.

In applying the results obtained from the gain survey to the larger population of subjects in the demonstration flights it was to be seen whether general relationships had been discovered which could be transferred to other users, who would then also show high levels of performance. The results presented in Figure 12 show that such a transfer took place, the performance of S1 being approached immediately and consistently by a majority of the company and visiting pilots, S3-41. It was thus evidently possible by a series of controlled runs with one subject to arrive at a reasonably general definition of optimum control conditions.

Comparison of Human and Automatic Operation

The results presented in Figure 12 also show a high level of success in meeting the requirements for Category II operation and thus suggest comparison with corresponding results obtained during automatic operation. For this purpose, and for correlation with previous work, it is desirable to establish a connection between different methods of error measurement. The present experimental method gives a value which is representative of the error during the course of a particular approach. Alternative methods, such as the one used by Morrall[6], give a value which is representative of the error at a particular height during a set of approaches. These different methods of dealing with the tracking error give the same result if a stationary random process exists, and on this assumption the present experimental results may be used to estimate the height error at a selected point, such as the Category II decision height of 100 feet.

From the present results, the time average of the vertical tracking error during an approach in smooth air was found to be 0.042 deg. This value was extracted as the mean of the integrated errors for the 36 visiting pilots, most of whom made three approaches. The corresponding one-sigma value was 0.053 deg for a normal distribution, or a height error of 1.8 feet at 100 feet. Other experimental results, with S2 as subject, indicated that this value should be doubled to allow for rough air effects, yielding a value which would be less than 3.6 feet for all types of air. A comparable value was found for automatic approaches in a similar (DC-9-30) airplane from the records of thirteen approaches at Oakland made with the final settings resulting from a gain survey[7]. Displacements from the glide slope were measured at a height of 100 feet, giving a standard deviation of 20.9 microamperes (based on a sample of N-1 = 12). The corresponding one-sigma value of the height error at 100 feet was 3.3 feet, using a conversion factor of 250 microamperes per degree for a 2-1/2 deg glide slope. Agreement between the results for HUD and automatic approaches was thus fairly close.

A similar equivalence is to be found in the results obtained previously by Morrall[6]. In his work, the one-sigma value of the height error at 100 feet was about 5 feet in automatic approaches and in approaches flown manually, using essentially the same head-up display format as in Figure 1. The vehicle in this case was a slow transport airplane, which may account for errors being somewhat larger than in the present work. On the other hand, since height errors were measured by ground theodolite they would include the effects of beam errors (although such effects are believed to have been small for the ILS system used). This comparison of human and automatic tracking performance is summarized in Table III which includes the results of Morrall, together with a result previously obtained with the format of Figure I in a jet fighter[8], in which only manual flight was possible. This table shows the similar levels for manual flight in jet aircraft, and equivalent tracking accuracy for manual and automatic operation in two kinds of aircraft (with a particular form of statistical equivalence assumed in one case).

The comparison can be extended by including results which are to be discussed in detail elsewhere[2]. Table IV shows how a conventional manual approach compares with an automatic approach, and with a head-up manual approach in which the relevant information is given optimum graphical form and dynamic characteristics best fitting the human operator. The first column shows how the systems compare for accuracy of flight. Conventional instrument flight is generally considered suitable for Category I conditions, with a possible extension to Category II conditions. Automatic operation is suitable for Category II or Category III conditions[9] and a similar accuracy has been demonstrated for HUD. The second column compares workload for the three systems, which is usually considered too high for comfort in any but Category I conditions of manual flight, while workload is moderate (perhaps low) for HUD[2] and, of course, it is low in automatic operation. In the third column, the available information is compared for the three systems. The completeness attributed to head-up operation is based on evidence of the pilot being able to move from the field of instrument information to the external visual field without interruption[2], a facility which is not available in conventional manual, or in automatic operation. In the fourth column, the possibilities of disorientation are compared. It is well known that the human pilot may become disoriented during a difficult transition between instrument and visual flight, while there is no such possibility in automatic operation. On the other

hand, there appears to be little likelihood of disorientation for the pilot using HUD, who acquires visual information as it becomes available in the external world, even if the display is out of alignment with the runway.

The comparison clearly shows advantages in all characteristics for HUD, and in most characteristics for an automatic system. The difference between these systems, however, is less obvious. Taking both accuracy and disorientation characteristics as approximately equal, for the two systems, the trade-off is between workload and completeness of information. The question is whether it is better to work a little and be well informed with HUD, or to work hardly at all and remain ignorant of factors affecting the total situation when an autopilot is used. For it is generally true that more information is connected with a safe approach than can be dealt with by an automatic system. The possibility of unwarranted intrusion into the landing area, or even of illegal activities, does not appear negligible. If such events can happen without the knowledge of the automatic system or the head-down monitoring pilot, or if they are perceived by the First Officer too late for effective communications, it may well be that the automatic system is not equipped to deal with all probable eventualities.

This inadequacy of an automatic approach system may prove unacceptable to the pilot, the man charged with responsibility for safe arrival of a passenger-carrying airplane. He may not accept an automatic system simply because it has been approved, when it has deficiencies lying outside the area of that approval. He may find difficulty in subordinating himself to a machine of less total capability than himself. And besides suffering from an isolation from relevant information, the automatic system is deficient in its provisions for announcing the possibility of an approaching and, as yet, undetected failure. The human pilot may therefore ask to be convinced that he could take over control under difficult conditions, at short notice. His ability to do so would depend on being able to manage all information bearing on the situation, continuously. This would include external world information since the subsequent manual control action might need to be assessed in relation to the aircraft's environment, especially for visual landings. These considerations suggest that an automatic approach should be monitored in the head-up mode, so that the pilot could continuously monitor the autopilot, know the status of the automatic system, and obtain such external world information as might be available. His overall information problem would then be solved, and since HUD can be flown to touchdown, he would be able when required to step into a situation which he could control.

The comparison made in Table IV does not include the probability of failure as a characteristic to be considered in judging between alternate systems because of the conjectural nature of the subject. Since the total number of landings made with either HUD or an automatic system is well below the total of ten million danger-free landings generally considered mandatory as a reliability target, the chances of achieving sufficient reliability are purely theoretical, and proponents of both systems are free to arrange contributory component failure rates to match the overall reliability requirement. It may be noted in passing, however, that the reliability of an automatic system should be computed after taking account of the chances that the flight instrument system would survive failure of the autopilot. Neither has the integrity of guidance information been discussed because the systems show equivalent tracking performances, so that defective guiding information, such as would result from a bent beam, would have the same effect on each system. On

the other hand, detailed attention has been given to human factors affecting the all-weather approach, showing that far-reaching effects are possible by designing an information system to suit the man, rather than by accepting only what happens to fit an airplane. By the systematic elimination of differences between the pilot's two main sources of visual information it becomes possible to avoid many shortcomings of conventional instrument flight. It is possible to achieve a method of operation in which there are less mistakes, whether of tracking or of judgment, through attention to the content, graphical form, position,and dynamics of the information presented. Perhaps the most important result is thus to undermine much of the case for automatic operation, which can no longer be justified as performing a service beyond human capabilities.

REFERENCES

(1) J. M. Naish, "Properties and Design of the Head-Up Display (HUD)", Report MDC-J1409, Douglas Aircraft Company, Long Beach, 1970 (previously Douglas Paper 4951, 1968).

(2) J. M. Naish, "Flight Tests of the Head-Up Display (HUD) in DC-9-20 Ship 382, November 1968 to January 1969," in preparation, Section 2.

(3) Reference (2), Section 3.

(4) Reference (1), Experiment 6.

(5) "Criteria for Approval of Category II Landing Weather Minima," FAA Advisory Circular 120-20, Appendix I, c, 3(c).

(6) J. C. Morrall, "Pilot's Safety Problem in Category II Operations and Potential Contribution of Head-Up Display," Royal Aircraft Establishment, Technical Report 66195, 1966.

(7) Anon., "Flight Development Program for a Category III All-Weather Landing System on the Model DC-9 Series 30," Douglas Report DEV-3795, 1968, Appendix A.

(8) J. M. Naish & R. Shiel, "Flight Trials of the Head-Up Display in Meteor and Hunter Aircraft," Royal Aircraft Establishment, Technical Report 65254, 1965, 75.

(9) O. B. St. John, "All-Weather Landing," Shell Aviation News, No. 364, 1968, 2-11.

TABLE II

Optimum Gains for DC-9-20 Aircraft found by Objective (O) and Subjective (S) Methods in Two Experimental Modes and used in Demonstration Flights

Mode	Method	Subject	1 K_ψ	2 $K_{\dot\phi}/K_\phi$	3 K_θ	4 $K_{\dot\theta}/K_\theta$	5 K_ψ/K_ϕ	6 K_{σ_E}/K_θ	7 K_{σ_A}/K_ψ
Aircraft	O	S1	i	1.25	0.3	1.0	i	(S2) 17	40
	S		0.075 (?)	1.0	0.2	1.4	i	15	i
Simulator	O	S1	0.15	1.5	0.3	1.0	1.2	(S1) 17	40
	S		0.125	1.25	0.2	0.8	1.2	15	i
Flight Demonstrations	-	S3-5	0.088	0.33	0.141	0.33	1.3	12.9 to 31	33.6
	-	S6-41	0.079	0.33	0.127	0.33	1.3	10.3 to 31	25.6

i signifies invariance

TABLE III

Comparison of Automatic and Manual HUD Tracking Accuracy:

Standard Deviation of Height Error at 100 Feet (Values in feet)

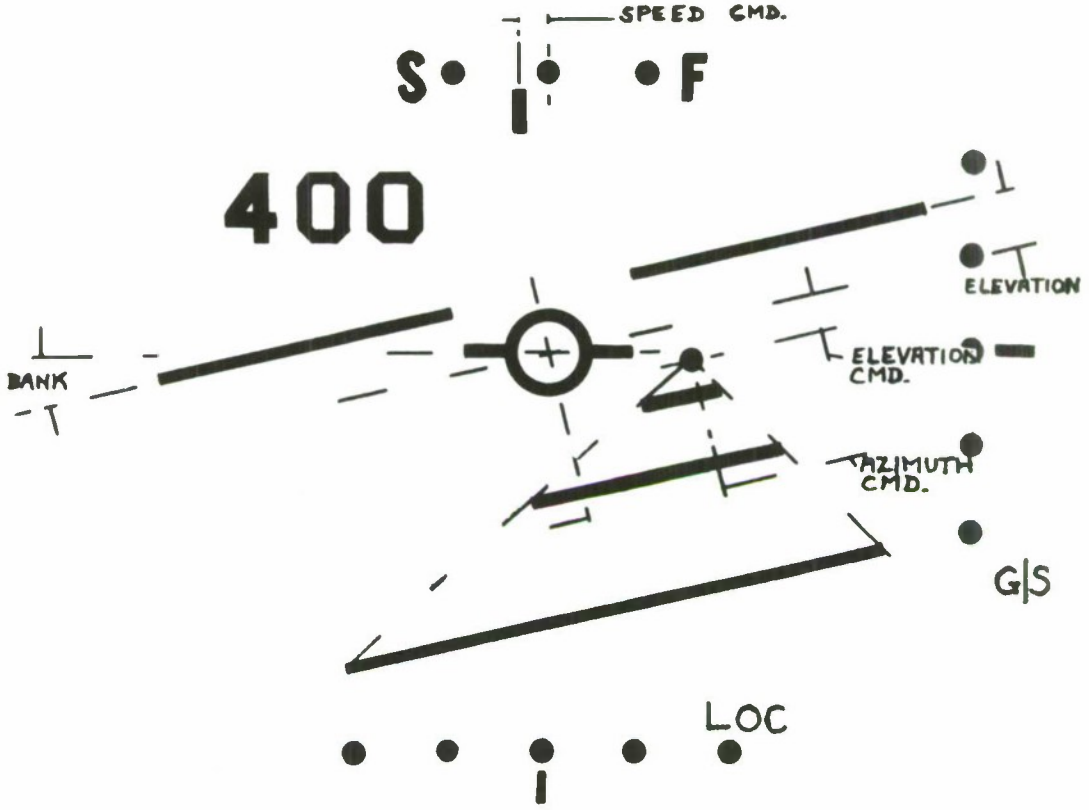
Vehicle	Slow Transport[6]	Jet Fighter	DC-9-20	DC-9-30
Automatic Approach	5	-	-	3.3
Manual HUD Approach	5	3.5	< 3.6	-

[6] Derived from Morrall, reference 6.

TABLE IV

Comparison of Manual, HUD and Automatic Approach Systems

Approach System	Accuracy	Workload	Information	Disorient-ation
Manual	Cat I - II	High	Incomplete	Possible
HUD	Cat II-III	Moderate	Complete	Unlikely
Automatic	Cat II-III	Low	Incomplete	Impossible



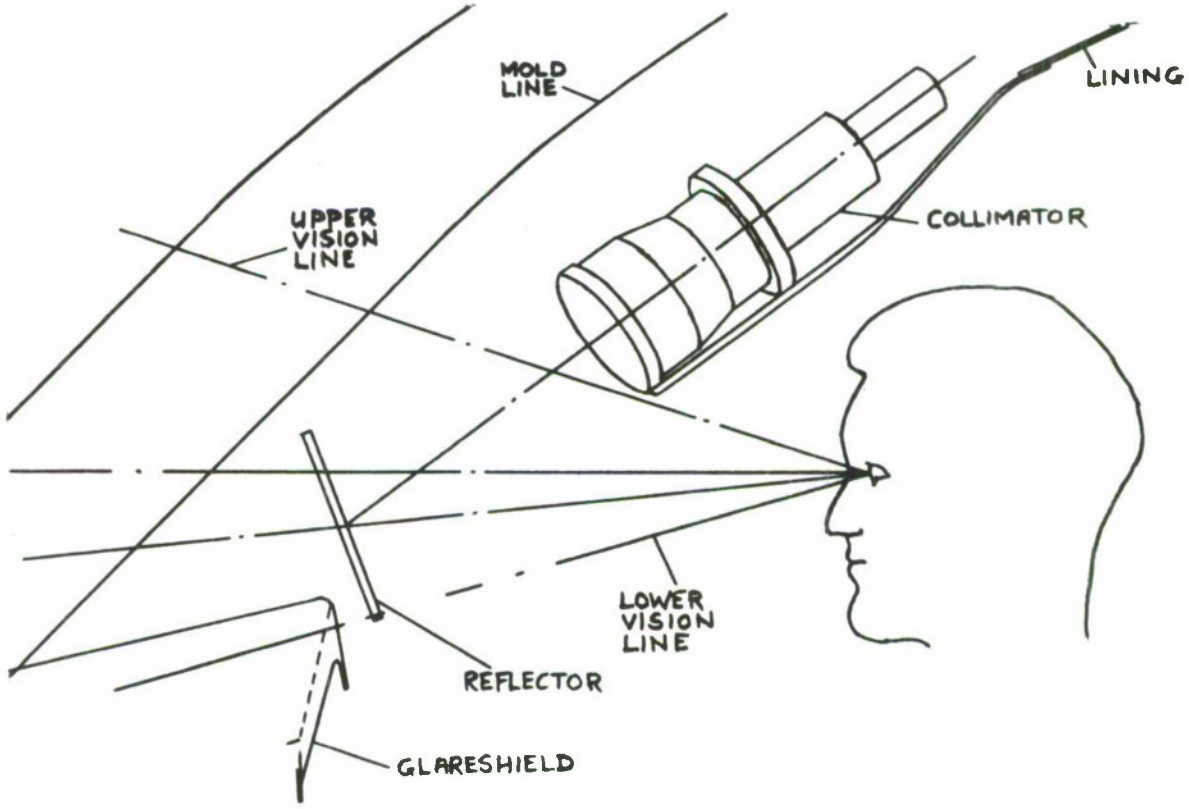
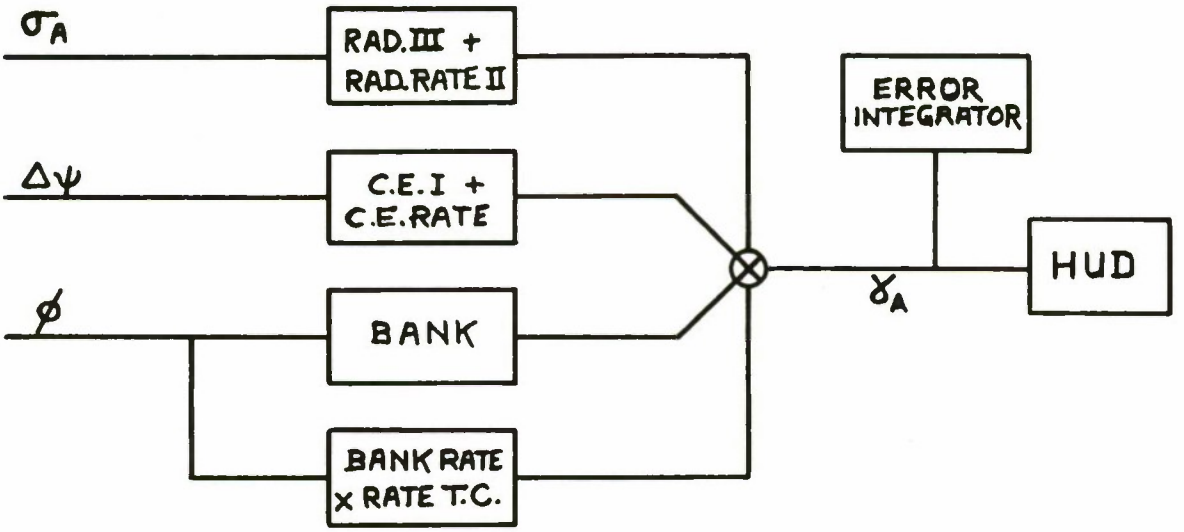


FIGURE 1 SYMBOL FORMAT

FIGURE 2 OVERHEAD COLLIMATOR INSTALLATION



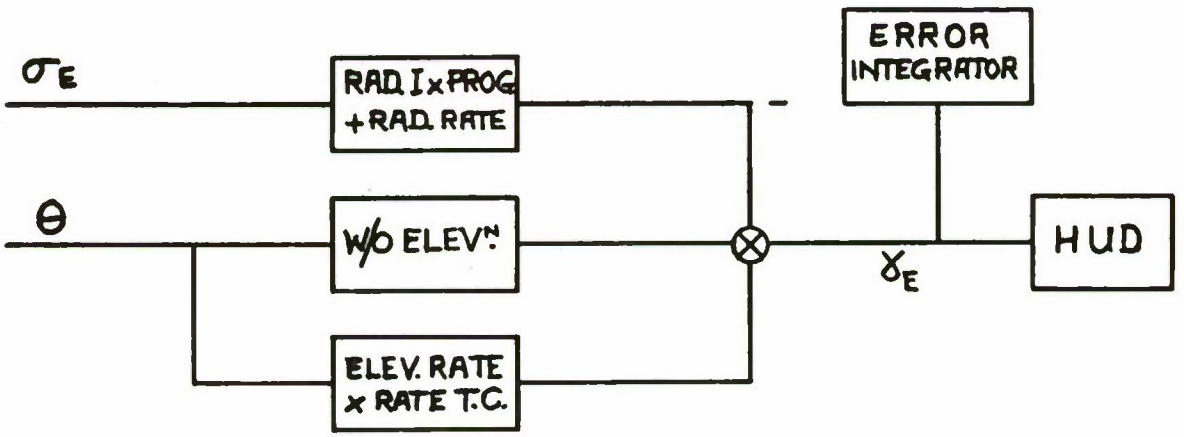


FIGURE 3 AZIMUTH COMMAND COMPUTATION

FIGURE 4 ELEVATION COMMAND COMPUTATION

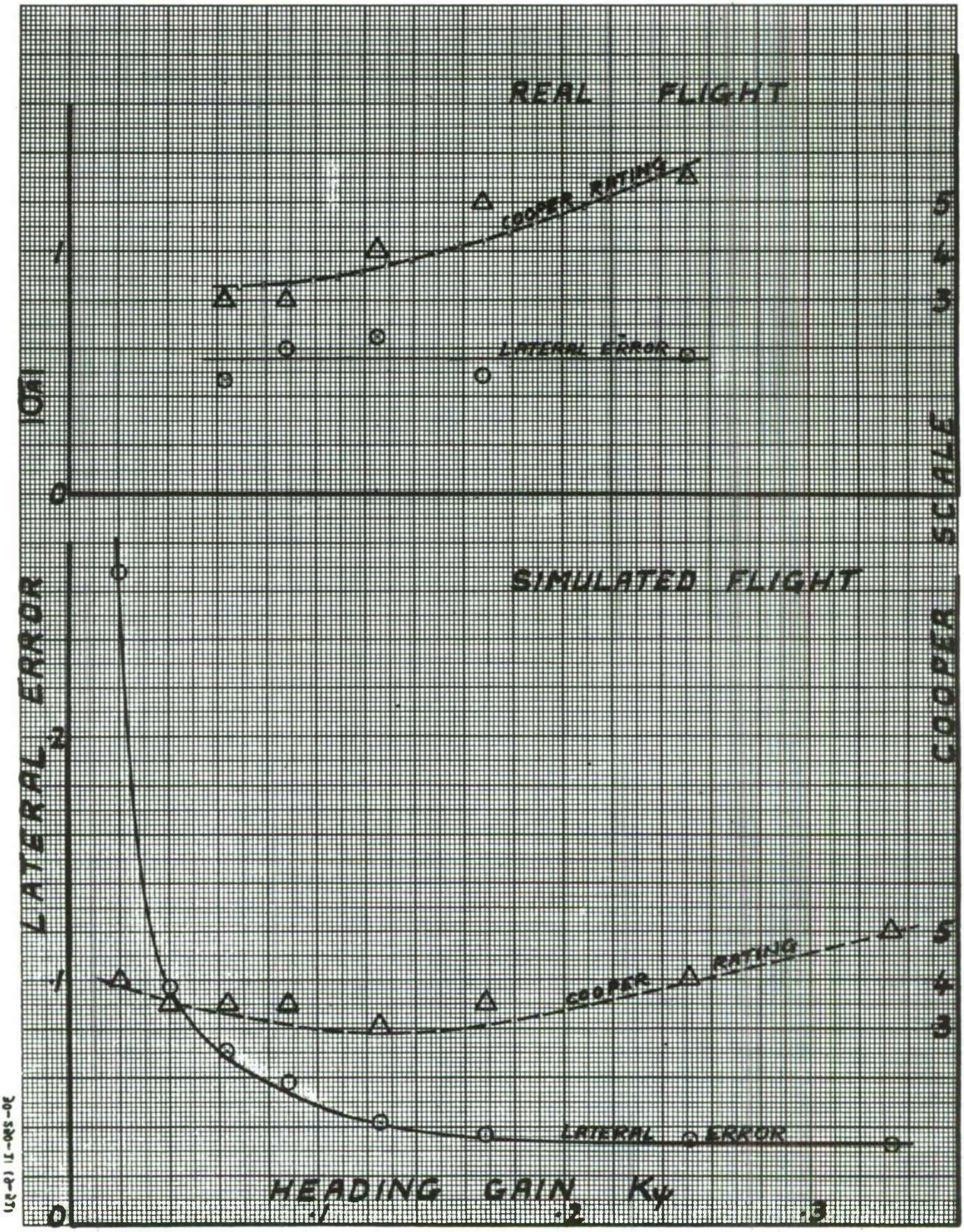

REAL FLIGHT
COOPER RATING
LATERAL ERROR
1
0
|σ|
3 4 5
COOPER SCALE
SIMULATED FLIGHT
LATERAL ERROR
-1
-2
COOPER RATING
5
3 4
LATERAL ERROR
HEADING GAIN Kψ
0
.1
.2
.3

FIGURE 5

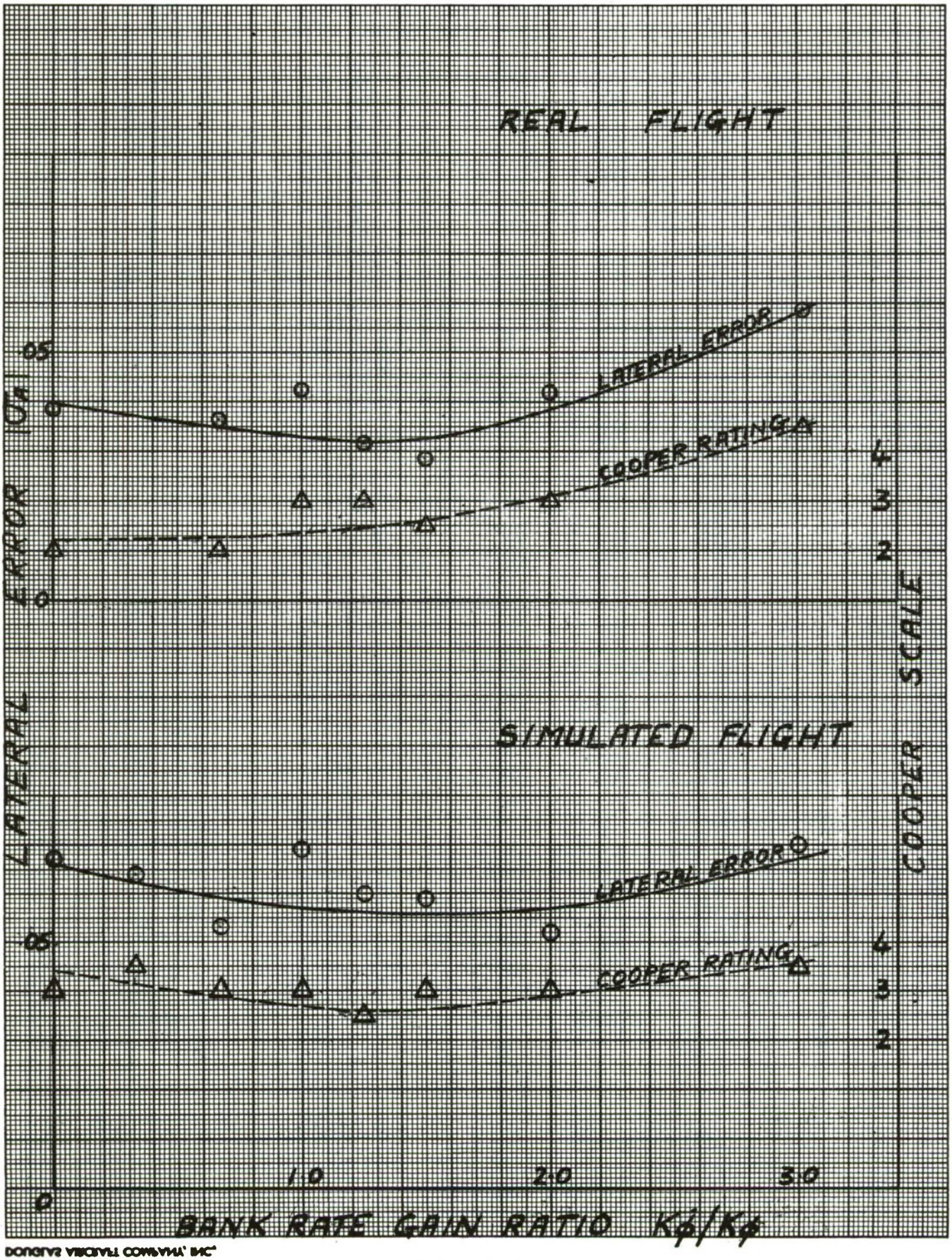

REAL FLIGHT
LATERAL ERROR
COOPER RATING
SIMULATED FLIGHT
LATERAL ERROR
COOPER RATING
LATERAL ERROR
.05
0
.05
0
1.0
2.0
3.0
4
3
2
4
3
2
COOPER SCALE
BANK RATE GAIN RATIO Kφ/Kφ

FIGURE 6

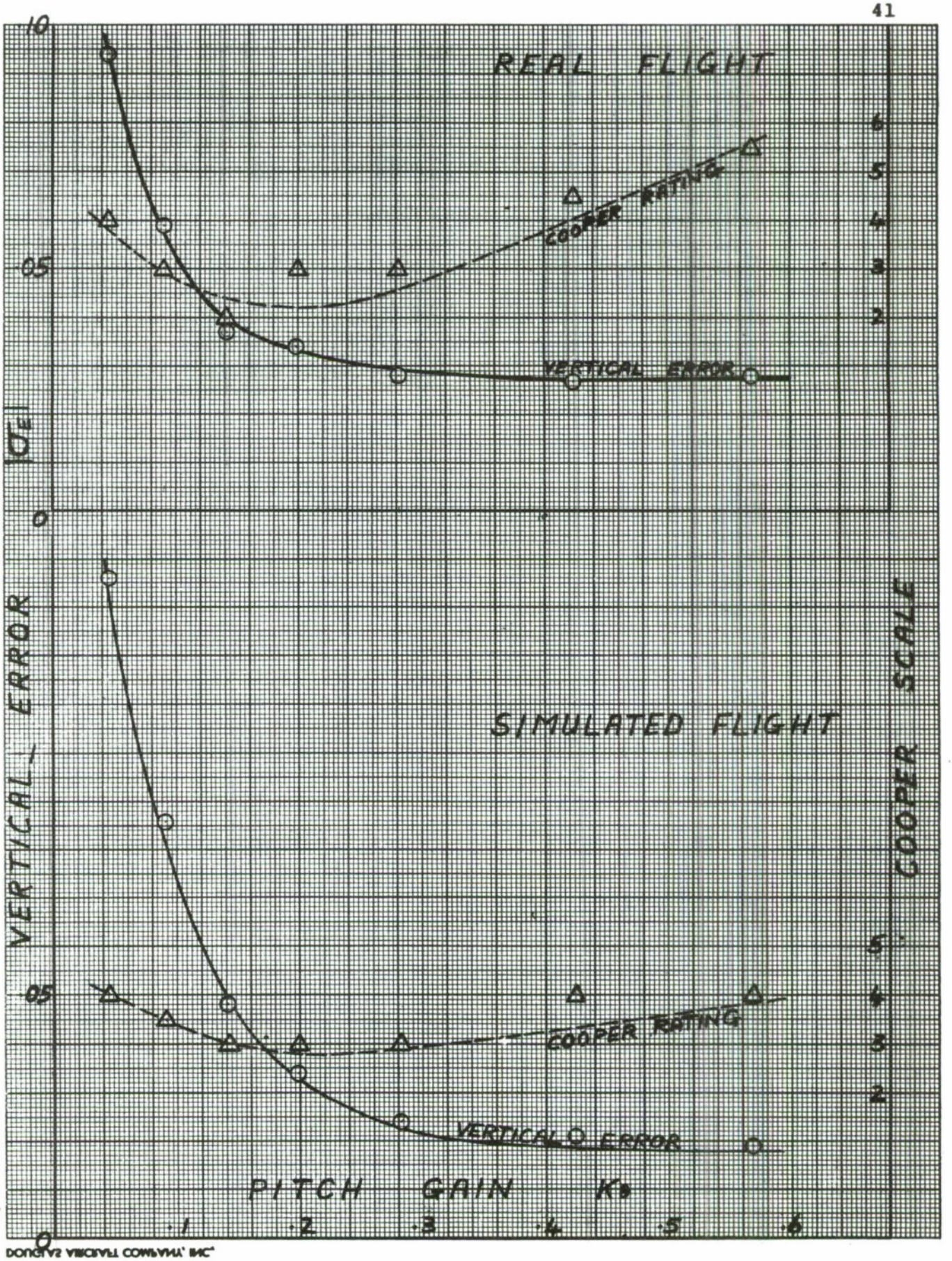

REAL FLIGHT
COOPER RATING
VERTICAL ERROR
SIMULATED FLIGHT
COOPER RATING
VERTICAL ERROR
VERTICAL ERROR |σE|
COOPER SCALE
PITCH GAIN Kθ
10
05
0
.1
.2
.3
.4
.5
.6
1
2
3
4
5
6

FIGURE 7

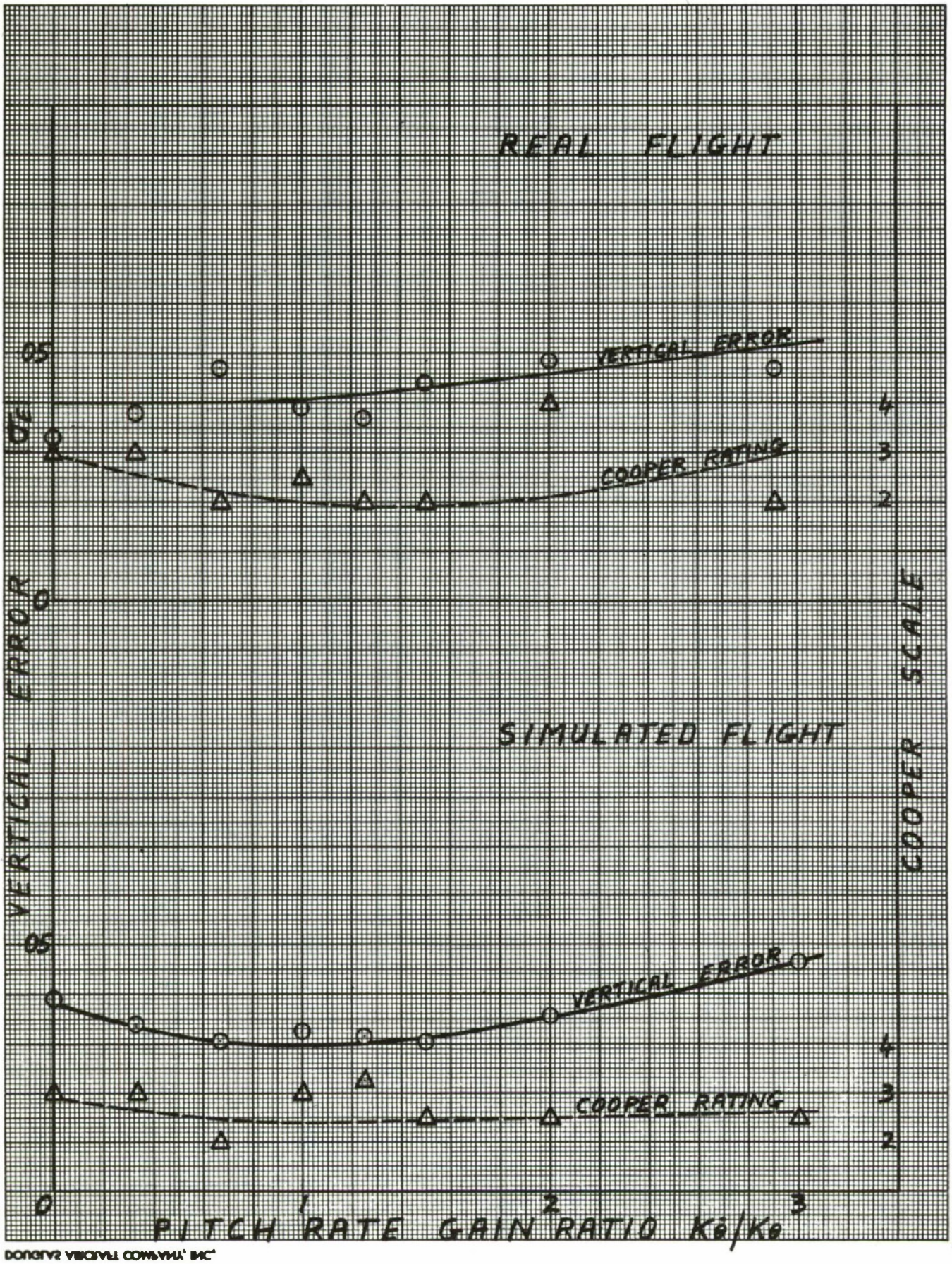

REAL FLIGHT
VERTICAL ERROR
COOPER RATING
.5
0
SIMULATED FLIGHT
VERTICAL ERROR
COOPER RATING
.5
0
1
2
3
4
3
2
VERTICAL ERROR
COOPER SCALE
PITCH RATE GAIN RATIO Kθ/Kθ

FIGURE 8

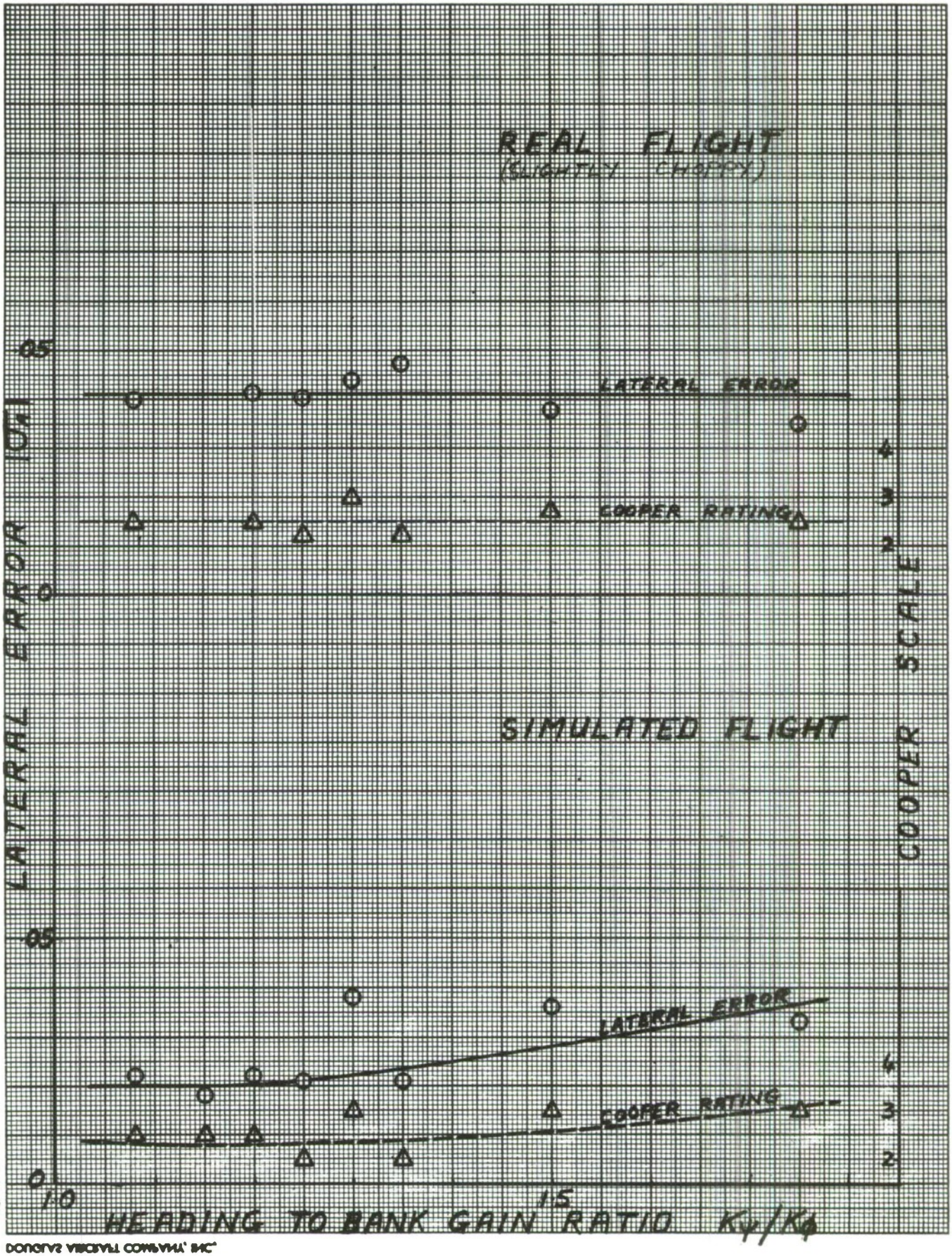

REAL FLIGHT
(SLIGHTLY CHOPPY)
LATERAL ERROR
COOPER RATING
SIMULATED FLIGHT
LATERAL ERROR
COOPER RATING
LATERAL ERROR [DA]
COOPER SCALE
.05
0
.05
0
1.0
1.5
HEADING TO BANK GAIN RATIO Kψ/Kφ
4
3
2
4
3
2

FIGURE 9

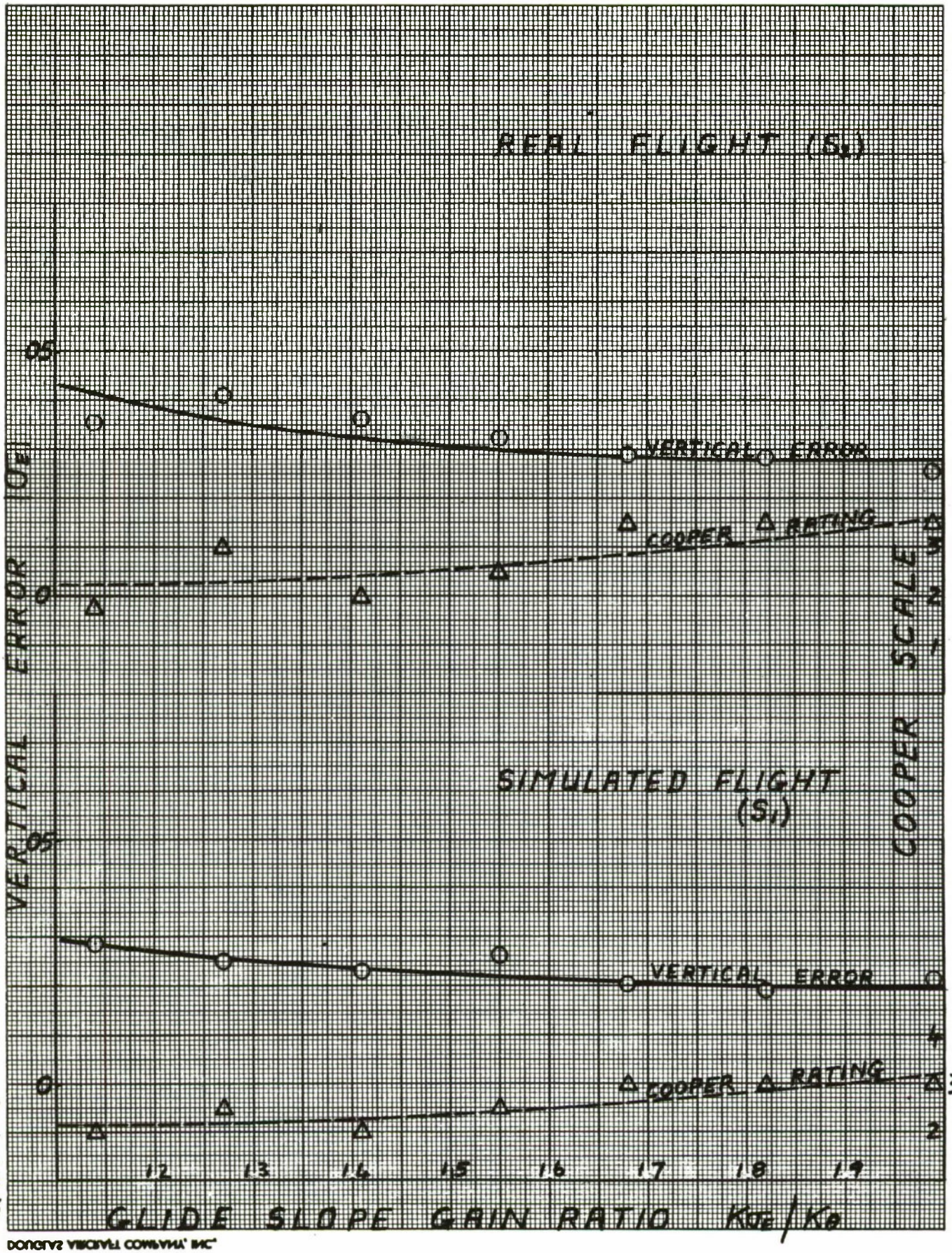

REAL FLIGHT (S_2)
VERTICAL ERROR
COOPER △ RATING
SIMULATED FLIGHT (S_1)
VERTICAL ERROR
COOPER △ RATING
VERTICAL ERROR 10_{E1}
COOPER SCALE
.05
0
.05
0
1
2
3
4
1.2
1.3
1.4
1.5
1.6
1.7
1.8
1.9
GLIDE SLOPE GAIN RATIO $K_{\theta E}/K_\theta$

FIGURE 10

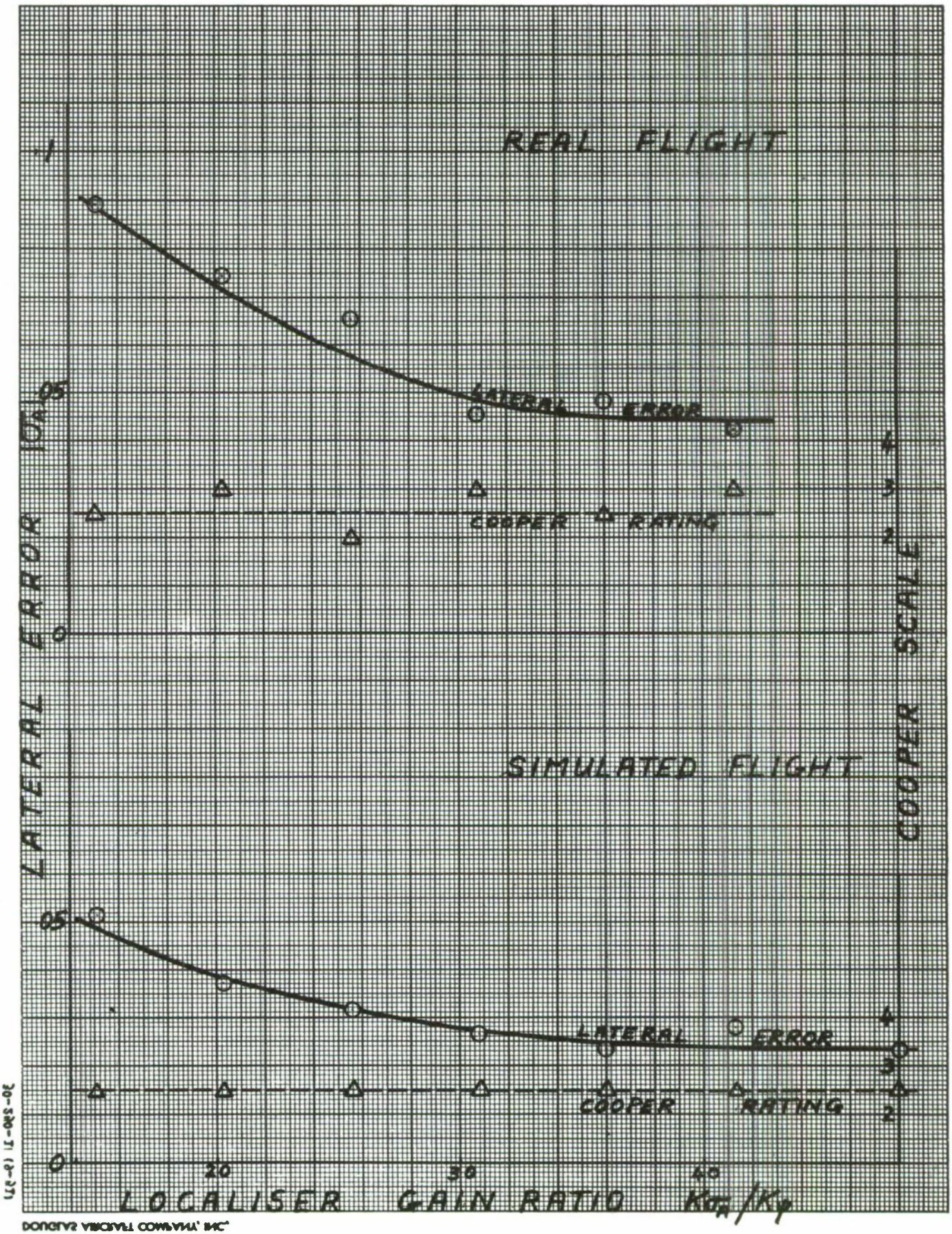

REAL FLIGHT
LATERAL ERROR
COOPER RATING
SIMULATED FLIGHT
LATERAL ERROR
COOPER RATING
LATERAL ERROR
COOPER SCALE
LOCALISER GAIN RATIO

FIGURE 11

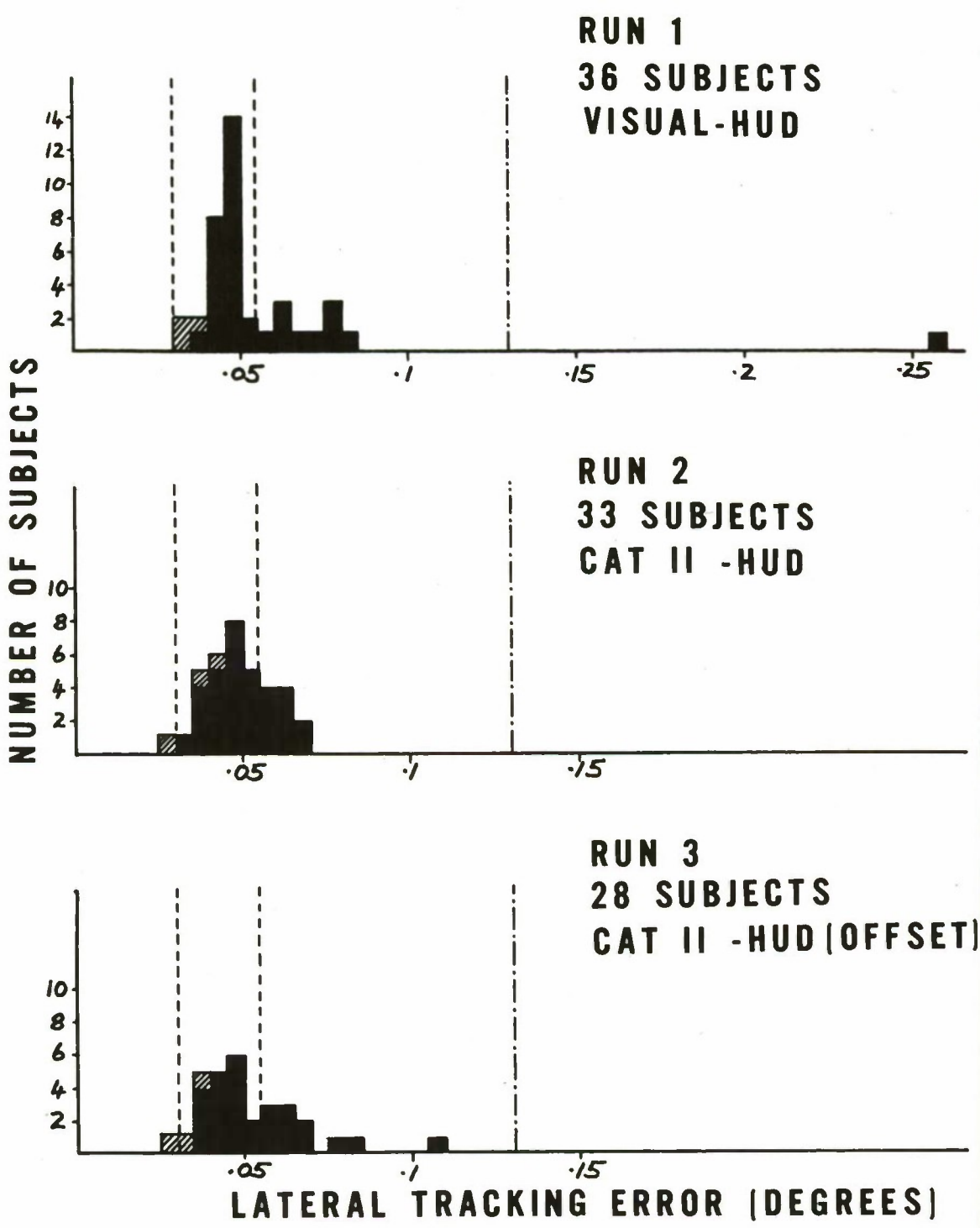


FIG 12
DISTRIBUTION OF LATERAL ERRORS FOR VISITING PILOTS

PILOT PERFORMANCE WITH A SIMULATED PICTORIAL LANDING DISPLAY INCLUDING DIFFERENT CONDITIONS OF RESOLUTION AND UPDATE RATE

By T. Wempe and E. Palmer
Ames Research Center, NASA
Moffett Field, Calif. 94035

SUMMARY

Four pilots participated in a study of the usefulness of a simulated pictorial runway display for making instrument approaches and touchdowns. Each pilot made 4 flights on each of 26 display conditions. The controlled variables were the resolution and update rate of the runway image. The display had good pilot acceptance and could be degraded by large variations in resolution and update interval without affecting pilot performance. However, the display had some deficiencies. During the approach, the display provided adequate lateral information but poor height control information relative to a three degree approach to the runway. During the touchdown the display was considered only marginal in providing information for control of distance down the runway, sink rate, and lateral displacement.

INTRODUCTION

Current trends in avionic development are toward providing an independent means of acquiring and displaying the information normally provided by the Instrument Landing System (ILS). A number of such systems have been proposed (refs. 1 and 2). Some use active transmitters on the ground, and others, such as radar or microwave receivers, may be purely passive. Two potential applications of these systems are to provide an independent pilot monitoring and manual backup capability for a fully automated approach and landing system, or to provide a system for low visibility manual approaches to airports not equipped with ILS transmitters.

The pilot's display for these systems is generally somc type of pictorial display of the airport runway plus some additional information in symbolic format (ref. 3). There are attributes to these kinds of displays: they provide an integrated and easily interpreted picture of the outside world, and other information pertinent to the piloting task can be easily added to the basic display. Inherent shortcomings of most of these display systems are that the picture of the airport that they produce is impoverished in content, degraded in resolution, and somewhat delayed in time.

The purpose of this study was to investigate the attributes and deficiencies of a pictorial runway display as a sole source of information for a pilot to perform a manual approach and touchdown. It was intended that the results of this study would (1) indicate thc feasibility of such a display, (2) indicate where improvements would be desirable (or necessary), and (3) provide baseline data for comparison with subsequent modified or augmented displays.

A small but proficient sample of professional pilots participated in three experiments. The objectives of Experiment I were: one, assess the pilot's ability to utilize a perspective display of an airport runway and horizon bar without any other aids in making simulated manual instrument approaches and touchdowns; two, determine specifically what information was inadequately presented by the display; and three, determine the effects of the runway display variables - resolution and update interval - on pilot performance. The objective of Experiment II was to determine the effects on pilot performance of removing the horizon from the display. The objective of Experiment III was to determine the pilots thresholds in perceiving lateral and vertical displacement errors when viewing the perspective runway display at fixed distances from the runway.

SYMBOLS

g	acceleration due to gravity	ft/sec^2
p	roll rate	rad/sec
q	pitch rate	rad/sec
r	yaw rate	rad/sec
s	Laplace operator	rad/sec
u	perturbed forward velocity	ft/sec
U_o	steady-state forward velocity	ft/sec
w	perturbed downward velocity	ft/sec
δ_α	aileron control surface deflection	rad
δ_e	elevator control surface deflection	rad
Θ	pitch angle	rad
ϕ	roll angle	rad
ψ	yaw angle	rad
$\dot{X}$	forward ground velocity	ft/sec
$\dot{Y}$	side ground velocity	ft/sec
$\dot{Z}$	vertical velocity	ft/sec
D	distance down runway from threshold	ft
L	lateral displacement from centerline	ft
S	sink rate	ft/sec
HR	horizontal resolution	deg
VR	vertical resolution	deg
UI	update interval	sec

TESTS AND PROCEDURES

Test Setup

To minimize the interaction of aircraft dynamics with the piloting task, a light aircraft was simulated for Experiments I and II. To reduce the effects of differences in piloting techniques, pilot options such as power adjustments and flap settings were not included in the simulation. In addition no instrumentation other than the horizon and the runway image was provided during data runs (an altimeter and a rate of climb were provided on the first practice day only) so that the pilot was forced to control the aircraft solely by reference to the pictorial display.

The piloting tasks of Experiments I and II and the judgement task of Experiment III were performed by pilot-subjects while they were seated inside a small portable cab. The cab was completely enclosed, and two fans provided a noise background to mask extraneous laboratory noises. A small fluorescent light mounted behind the pilot's backrest provided low level indirect lighting. The pilot had three controls: a side arm controller with a flexible fiber control stick, a start button to start the flight, and an abort button to signal a desire to abort a landing. Figure 1 depicts the interconnection between the cab and the simulation equipment.

Vehicle dynamics - For Experiments I and II the dynamics of a Navion, a low-wing four passenger light plane, were simulated on a digital computer using the equations described in Appendix A. The power and trim were set to maintain an airspeed of 176 feet per second (104 knots) when on a 3° glide slope. No throttle was provided but the aircraft could maintain level flight with some loss in airspeed and resulting change in trim. The yaw rate was computed so that the aircraft always made coordinated turns, hence rudder control was not required or provided.

For Experiment III, position judgments, the display perspective program was retained but the aircraft dynamic program was replaced with a program which allowed the pilot to vary his altitude or lateral displacement at a fixed range from the runway threshold. A constant control stick displacement caused a proportional rate of change of position in the appropriate direction.

Flight conditions - For each flight of Experiment I, the simulated aircraft was positioned 10,000 feet from the threshold on a 3° glideslope to a point 1,000 feet down the runway (fig. 2). At the beginning of the flight, the miniature aircraft symbol in the display (fig. 3) was centered on and level with the horizon. The aircraft was trimmed such that, if there were no wind conditions, it would proceed down the proper flight path providing the initial position of the miniature aircraft was perfectly maintained by the pilot. If wind conditions were present, appropriate adjustments in pitch and heading were required to maintain the proper flight path. Flight conditions for Experiment II were the same as Experiment I except that the horizon was removed.

Disturbance - For Experiments I and II constant crosswinds of 0, ±3, ±6 or ±9 feet per second and constant vertical drafts of 0, ±1.5 or ±3 feet per second were entered into the aircraft dynamic program and sustained for the first half of the approach. Turbulence was not simulated.

Display - The display, fig. 3, was generated by a digital computer and was a perspective view of the runway depicted in fig. 2. The field of view of the perspective display was 40° by 40° in the "real world." In the vertical direction 7° was visible above the aircraft centerline and 33° below the centerline. The display was generated on a cathode ray tube and transmitted by a 525 line closed-circuit TV system to a monitor in the pilot's cab. The pilots viewed the TV monitor from a distance of about 18 inches. The display was thus compressed to a scaling of approximately 1:2 as compared to a "real world" view of the runway.

The resolution of the displayed runway image was degraded by selectivity reducing the effective horizontal and vertical digital resolution in the computer display generation program. The update interval or the time between new frames of information, was simulated by adjusting the updating in the digital program and could be varied in increments of .05 second. When present the horizon and instruments always were updated every .05 second. The horizon always had a resolution of .05° relative to the "real world." All display elements were refreshed at the rate of 60 frames per second to prevent flicker.

Test Subjects

Four pilots with extensive backgrounds in instrument flying were selected to participate in these experiments. They had very little prior experience with pictorial displays or laboratory simulation experiments. All had recent experience in flying light aircraft. Their experience is summarized in the following table.

Pilot	Experience	Estimated Flight Time (Hours) Instrument	Night	Total
1	General aviation-commercial instrument flight instructor	1,200	2,000	15,000
2	General aviation-commercial instrument flight instructor	180	336	3,306
3	Heavy piston aircraft military pilot	500	1,000	10,000
4	Heavy jet aircraft airline pilot	2,000	15,000	30,000

Procedure

Instructions to pilots - Prior to each experiment the pilots were provided with written instructions which described the purpose of the experiment and the specific tasks they were to perform. These instructions are included as Appendix B. Some items in the written instructions were reinforced verbally and all questions were answered; however, no information was offered as to how to detect glide slope and localizer errors from visual cues provided by the pictorial display.

Performance measures and pilot opinion - For each of the simulated flights of Experiments I and II, altitude, lateral displacement from an extension of the runway centerline and sink rate were recorded at distances from the threshold of 5,000, 1,000 and 0 feet. These same measures were also recorded at the instant of touchdown along with the additional measure of the distance down the runway from the threshold. If the pilot pressed the abort button during a flight, the aircraft altitude and distance from the threshold were recorded at that point. At the end of each set of four flights for a given display condition, each pilot rated the usefulness of the display for both the approach and landing phase using the Cooper-Harper rating scale (ref. 4).

For Experiment III, position judgments, the lateral displacement or the height was recorded when the pilot pressed the button indicating that he had positioned the aircraft as instructed.

Training and experimental design, Experiment I - Each subject spent two hours a day for five days on Experiment I. The first day was used for general orientation to the project including instructions for Experiment I and practice in flying the simulation. On this day only, altitude and vertical velocity were displayed on the two instruments shown at the bottom of figure 3, and the pilots were told their touchdown performance after each landing. During this first day, various display conditions and wind conditions were gradually introduced into the task.

The 26 display conditions used in this experiment, figure 4, were divided into two groups having approximately the same difficulty relative to resolution and update interval. The sequence of each group was randomized and used as the sequence of display conditions for days two and three for pilots 1 and 3. A reverse order was used for pilots 2 and 4.

Each pilot made four flights at each display condition. The first flight had no wind. For each of the remaining three flights, a random sample without replacement was drawn from the crosswinds — 3, 6 and 9 feet per second — and from the drafts — 1.5, -3 and +3 feet per second. Except for the two largest drafts which were already given signs (-3 and +3 feet per second), the signs for the remaining winds were selected randomly. A positive wind was an updraft or a crosswind from the right.

Days two and three were considered practice days and not scored. The experimental conditions for days four and five were identical to days two and three and performance data were taken.

Training and experimental design, Experiment II - This experiment was conducted after the five days experience on Experiment I, and differed only in that the horizon was removed and fewer display conditions were used. After reviewing the instructions for Part II, only four flights for warm-up were allowed and then data flights were made. All pilots made two sets of four flights with each of the following two display conditions: (horizontal resolution = 0.1°, vertical resolution = 0.1°, update interval = 0.05 sec) and (HR = 0.4, VR = 0.8, UI = 0.30). In addition, pilots 2, 3 and 4 made two sets of four flights each of the two conditions: (HR = 0.1, VR = 0.1, UI = 1.0) and (HR = 0.4, VR = 0.8, UI = 1.0). The conditions for the sets of four flights each were presented in random order to each pilot. Winds for each set of four flights were selected in the same manner as that described in Experiment I.

Training and experimental design, Experiment III - Combinations of the two display conditions (HR = 0.1, VR = 0.1) or (HR = 0.4, VR = 0.8), along with two distances to the runway threshold, 1,000 feet or 5,000 feet, comprised the four experimental conditions of this experiment. For each of these four conditions, the pilot was required to make, first, eight vertical judgments for initial positions alternately above and below the three degree glideslope, and then, make four horizontal judgments for initial positions alternating between the right and the left of the runway centerline. While judgments were being made in height, the computer program maintained the aircraft on the runway centerline; and while judgments were being made relative to the runway centerline, the computer program maintained the aircraft on the three degree glideslope.

The experimental design was a Latin square of the four conditions balanced among the four pilots. The sequence for conducting this experiment was to first review the instructions with the pilot. Then a condition that was the same as the last condition in the Latin Square for each pilot was given for practice and not scored. Finally, the conditions were presented according to the Latin Square.

Pilot debriefing - At the end of Experiment III each pilot was debriefed. The results of the debriefing are given in Appendix C.

RESULTS

Though many statistics were computed from the performance measures and carefully reviewed, most of the information provided by them was redundant. Therefore, only those measures which are both easily understood and which retain the essence of the results are included in the following discussion.

Experiment I, Flights with Horizon on Display

Overall performance - The approach task without winds was principally an inner-loop control task (ref. 5), that is, for the pilots to control height and lateral position

when no winds were present they had only to minimize the pitch and roll errors caused by lack of perfect trim and pilot-controller noise. This could be accomplished by holding the miniature airplane symbol level on the horizon bar and centered laterally on the runway. Since this was the nominal position of the airplane symbol on the three degree glideslope, the simulated aircraft would then progress down the desired glide path to the runway with only moderate errors developing in height and lateral displacement.

Table 1 presents the height and lateral displacement statistics when the simulated aircraft was 1,000 feet from the threshold for conditions of no wind and the most severe winds. (As will be shown in the next section, there was no particular need to segregate performance with the different display conditions.) The no wind data in this table indicate that the typical height error from the nominal altitude of 105 feet at this distance was several times larger than the lateral position error. This suggests that the pilots were getting better lateral displacement information than height error information and were able to improve their performance in outer-loop lateral control beyond that achieved solely by minimizing the inner-loop errors as described above.

Compared to the no wind approaches, performance with winds present showed much higher errors relative to the desired flight path. At a distance of 5,000 feet from the runway threshold, which was the position where the winds were terminated, all of the pilots arrived in the vicinity of the altitudes that the drafts would have placed them if they had made no adjustments in pitch to compensate for the drafts. The data of table 1 indicate that none of the pilots had completely recovered from the effects of the drafts by the time that they were 1,000 feet from the threshold, with pilot number 3 having made only small corrections and pilot number 4 having made essentially no corrections.

Though there were some differences among the pilots in lateral control, all of them had made some lateral corrections for the crosswinds by the time they had reached a distance of 5,000 feet from the runway (pilots 1 and 2 had corrected for about one-half of the effect of the crosswinds and pilots 3 and 4 had corrected for about one-third of the effect of the crosswinds). By the time the flights had reached a distance of 1,000 feet from the threshold (table 1), only pilot number 2 showed a significantly non-zero average lateral displacement error (21 feet) in the direction blown by the crosswinds; however, all of the pilots showed significantly larger standard deviations (sigmas) in lateral position due to the crosswinds compared to those flights where no winds were present.

As with the no wind data, the height errors with winds present were several times larger than the lateral displacement errors, thus supporting the statement that lateral displacement information was much better than height error information with this pictorial display.

At this point it is evident that this type of display provided the pilots with inadequate height information. This agrees with the results of another study (ref. 6) in which pilots visually flew a simulated aircraft with no altimeter on a night approach from 18

miles to 4.5 miles to touch-down. The pilots were easily influenced into making gross errors in height estimation and glideslope control by tilting a simulated night scene of a city beyond the airport. Apparently a runway scene alone does not supply enough information for a pilot to judge height and glide path angle. This referenced study also noted a large variability among pilots in their ability to utilize a runway scene for glide path control.

A chi-square test of the touch-down performance for each pilot for all display conditions combined and for no crosswinds or vertical drafts (fig. 5(a)) indicated no significant differences in performance among the pilots ($p > 0.05$). Thus, with no winds all of the landings for the 104 no wind flights were on the runway with about 75% of them considered successful, 25% marginal and 0% unsuccessful. (Qualitative definitions for landing performance are given in table 2.) Of the marginal touchdowns, the relative causes for this classification were: sink rate - 75%, distance down the runway - 18%, both distance and sink rate - 7%, and lateral position - 0%.

When winds were present, there were significant differences in touchdown performance among the pilots. Figure 5(b) depicts the touchdown performance for all display and wind conditions for each pilot. A chi-square test of the data used for this figure indicated that the pilots differed very significantly in performance ($p < 0.01$). In fact, any two of these pilots were significantly different from each other.

Only pilot number 2 performed as well at touchdown when the winds were present as he did when they were not; the other three pilots showed varying decrements in landing performance, which, in almost all cases, stemmed from their inability to recover from height errors introduced by up or down drafts. With the exclusion of 9 unsuccessful landings (off the side of the runway or sink rate greater than 8 feet per second) by pilot 4, all unsuccessful landings for all pilots were either short of the runway or off the end of the runway and were directly attributable to height control errors caused by the drafts.

To obtain an estimate of what touchdown performance might have been like if the display had been augmented to improve height control, touchdown performances for the 104 flights for all pilots with no winds present were examined. For these flights the lateral displacements at touchdown had a mean of 1.6 feet and a 2σ value of 21.2 feet. For comparison, the FAA specifications for a fully automatic landing system (ref. 7) states that the two-sigma value of lateral displacement must be less than 27 feet with no touchdowns occurring within 3 feet of the runway edge. In this respect, the data of this study are considered only marginally satisfactory (the simulated Navion was fairly easy to control laterally), and the display should be improved to allow more precise lateral guidance near touchdown. The distances down the runway for flights with no winds present had a mean of 1,369 feet and a two-sigma value of 1,037 feet. For this performance measure, the FAA specified a 2σ limit of 1,500 feet and that all touchdowns must occur 300 feet beyond the threshold for a fully automated system. Though the 2σ value of the data of this study was within the FAA specification cited, it still seems excessive for a light plane (even without throttle control). It was thought that

improved approach height control might reduce the dispersion of touchdown distance; however, an investigation of the correlation between height at 1,000 feet from the threshold and touchdown distance indicated otherwise. For pilots 1 through 4, the product-moment correlations between height and distance for no wind flights were $r = 0.39, 0.78, 0.28$ and $.00$, respectively. For all pilots together, this correlation was 0.29, indicating that only a 5% reduction in overall distance dispersion could be achieved by rigorously controlling height. More study is required relative to improving this performance measure.

The sink rates at touchdown for the no wind flights (which, as mentioned above, accounted for 75% of the marginal touchdowns) had a mean of 4.0 feet per second and a 2σ value of 2.7 feet per second. This is considered excessive for the aircraft simulated and suggests that the runway outline alone is inadequate in providing height or sink rate information when near touchdown. Conceivably, this information could be obtained from some other source and incorporated in the display to improve performance.

Effects of display variables - The effect of display resolution and update interval on landing performance is presented in figure 6. For example, figure 6(a) presents the touchdown performance for the displays having horizontal resolutions varying from $0.1°$ to $1.6°$ while the vertical resolution and update interval were held constant at $0.1°$ and 0.05 seconds, respectively. The performance of pilot number 4 was highly related to the draft condition for all display conditions, and to prevent obscuring any small effects of the controlled variables, his performance data were excluded from this figure only. It is apparent from this figure that there was no consistent or sizable effect on performance due to the display variables except for the extreme conditions of resolution or update interval. These conclusions were also confirmed by analyzing the means and sigmas of the height and lateral offset data taken at 5,000, 1,000 and 0 feet from the threshold as well as the three touchdown measures, sink rate, distance down the runway and lateral position.

From these performance data, it appears that if the horizontal resolution were as good as $0.4°$ and the vertical resolution were as good as $0.8°$, while the update interval was as fast as 0.5 seconds, then the performance as measured in this experiment would not be substantially improved by further improvement in the quality of this particular display.

Pilot ratings of display conditions - Figures 7(a) through 7(f) show the results of the pilot ratings for the approach phase and the landing phase of each series of flights at a given display condition, averaged over the four pilots. In general, the pilots rated the type of pictorial display used in this study as slightly more useful for the approach task than for the landing task, though none of the display conditions was rated satisfactory. Though the pilot's range of ratings (highest minus lowest rating) for each display condition and flight phase varied as much as zero for one particular condition to six for another, the variability appeared random among the display conditions; hence, the average range of 2.5 Cooper-Harper units is used to describe the variability of the ratings among the four pilots for any display condition.

Figures 7(a) through 7(c), the results from variations of one-at-a-time of each of the three controlled variables about the best display condition (HR = 0.1, VR = 0.1, UI = 0.05), suggest that for these pilots there were regions of indifference. When the vertical resolution was 0.4° or less, or the update interval was 0.3 seconds or less, the ratings for the landing phase were almost constant at a rating of 4.5. Above these values of the controlled variables, the corresponding ratings were higher. Though the ratings varied only slightly for horizontal resolutions of 0.1° through 0.4°, the pilots tended to prefer the best resolution presented in this direction. For the approach phase, the ratings were similar except that they were lower by about 0.5 units and the indifference to increases in the update interval extended to 0.5 seconds.

When the controlled variables were changed one-at-a-time about the display condition (HR = 0.4, VR = 0.8, UI = 0.3), the resulting ratings were higher and less consistent (figs. 7(d) through 7(f)). The apparent region of indifference, though less defined, increased to about 0.8 degrees of vertical resolution and 0.4° of horizontal resolution. With these values for resolution, the update interval appeared to have no consistent effect on the ratings.

Note that the regions of indifference discussed above are valid only for variations in one display variable at a time. When two or more display variables are changed concurrently, a higher rating might be obtained than would be expected from a simple combination of the one-dimensional effects.

When these ratings are compared with the performance data, it appears that the pilot ratings were more sensitive to changes in the display variables and preceded changes in performance. This result is typical of piloting tasks where workload or pilot insecurity with the task increase before performance degrades.

From these ratings and the performance data, a reasonable compromise in display quality can be stated. For the approach from 2 miles to 0.25 miles, a display that is as good as (HR = 0.4, VR = 0.4, UI = 0.5) is desirable, and for the landing, a display that is as good as (HR = 0.4, VR = 0.4, UI = 0.3) is desirable. In either case, additional information will have to be added to the display to make it acceptable.

Desire to abort the landing - The number of landings aborted as a function of the display variables was highly correlated with the pilot ratings. This indicates a consistency between pilot opinion and pilot reaction to each display condition.

Experiment II, Flights with No Horizon on Display

Figure 8 presents a comparison of landing performance with and without the horizon included in the display. For the best display condition (HR = 0.1, VR = 0.1, UI = 0.05) the runway pattern itself provided much of the necessary inner-loop information on pitch and roll, and performance was only slightly degraded without the horizon. As the display was made more coarse in resolution and slower in update interval, the loss of the inner-loop information had a dramatic impact on performance.

Table 3 compares the pilot ratings for each display condition presented both with and without the horizon. This table indicates that even for the best display (HR = 0.1, VR = 0.1, UI = 0.05), the pilot ratings were about one unit higher when the horizon was not displayed. As the display was degraded, the ratings for the no horizon display (particularly for the landing phase) increased considerably, indicating the lack of utility of such a display.

The results of Experiment II demonstrate the necessity of including a real-time horizon on pictorial displays that are less than perfect in resolution and on those which have delays in the presentation.

Experiment III, Position Judgments at Fixed Distances

Lateral position - Since there was no interest in right and left side differences in judgments, the signs of the lateral distances from the runway centerline for alternate judgments were changed and the data subjected to an analysis of variance (ANOVA). The pilots again differed significantly among themselves. However, the most important result was that there was no significant difference between the two levels of resolution — (HR = 0.1, VR = 0.1) and (HR = 0.4, VR = 0.8). At a distance of 1,000 feet from the threshold, the mean judgment for all pilots together was 9.0 feet to detect a displacement error and the sigma was 7.0 feet; at 5,000 feet, the mean was 10.3 feet and the sigma was 18.8 feet.

Note that the sigmas for lateral displacement performance in Experiment I for each pilot under no wind conditions (table 1) are only slightly larger than 7.0 feet, thus, showing a correspondence between visual perception as measured in Experiment III and performance in the lateral direction. When crosswinds were present in Experiment I, the variability in performance increased over the variability in the judgment data, but, with the exception of pilot 2 there were no statistically significant constant errors in lateral position. This suggests that the pilots were attempting to reduce the perceived error but showed an effect from the crosswind disturbance.

An interpretation of the judgment data is that 97% of the time it can be expected that these pilots would perceive a lateral error of 33 feet (mean plus 2σ) at a distance of 1,000 feet from the threshold or 48 feet at 5,000 feet from the threshold. These two displacement magnitudes, which are highly likely to be detected, are 0.38° and 0.31°, respectively, from the center of the localizer beam and indicate that the information for good lateral control was adequate.

Vertical position - Because of the asymmetry of the vertical judgment task, the judgments made from below the 3° glide slope were kept separate from those made from above. A summary of the results from an ANOVA is that (1) the pilots were different in performance at this task, (2) there was some pilot by range or by resolution interactions, (3) range had a small but significant effect on judgments measured in

degrees relative to the glide slope (the judgments from above or below at 5,000 feet were higher by 0.36° than those made at 1,000 feet), (4) there was a constant error in estimating the position of the 3° glide slope (average estimation was 3.44° at 1,000 feet and 3.80° at 5,000 feet), and (5) the judgments for the two conditions of resolution were not significantly different.

The data of figure 9, derived from the judgment statistics presented in table 4, clearly indicate the lack of utility of an unaided pictorial display in controlling height to an acceptable tolerance. Even if training were employed, it is unlikely that height judgments would be better than the two sigma value of 0.5° demonstrated by Navy carrier qualified pilots estimating glideslope angles to a carrier deck (ref. 8). With perception errors this large, desired performance within ±0.35° (FAA Category II specifications, ref. 9) of the established glide slope would be unobtainable.

After the foregoing, it is of only academic interest to note that the sigmas (in feet) for the pilots making vertical position judgments at the 1,000 foot distance were very similar to the sigmas for height performance with no winds (table 1) — with allowance made for the unexplainably high value for pilot 1. However, when the winds were present (table 1), it appears that the pilots concentrated more on the pitch control task than the height control task (in varying degree among pilots) and they incurred large constant height errors that should have been apparent, considering the judgment data.

Personal Interview

A summary of the pilots' comments on each question of the interview is given in Appendix C. In general, the pilots thought that the pictorial display was easier to use and incurred less workload than an ILS cross-pointer instrument. The pictorial display used in these experiments needs to provide additional or more precise information on (1) height, (2) sink rate, (3) distance progressed down the runway and (4) lateral offset when near touchdown. In addition, with long update intervals, heading information was not good.

The need for additional information on height and sink rate for pictorial displays was recognized in a study of an airborne television display, ref. 10, which supports the pilots' opinions on items (1) and (2) above.

Additional instrumentation requested by the pilots in order of desirability were: rate of climb, altimeter, airspeed, heading, ILS indicator, turn and bank, and power.

CONCLUDING REMARKS

An unaided pictorial runway display as a sole source of information in manually controlled, zero-visibility approaches and landings has been studied by simulation to determine its attributes and deficiencies. In general, the display had good pilot

acceptance and could tolerate large variations in resolution and update interval without affecting pilot performance; however, the display was inadequate in providing information for the judgment of glide slope errors and the control of height.

During the approach, the display provided good information about small lateral displacement errors and enabled good lateral control. On the other hand, during this phase of the flight, not only was the judgment of height relative to the glideslope poor, but also, the pilots tended to perform even more poorly in height control than their perception of height errors would indicate.

During the landing phase of the flights, the display was considered only marginal in providing information for the control of lateral displacement and distance down the runway at touchdown. Though the display had some utility in controlling sink rate at touchdown, it was not considered to be adequate for this purpose.

It is believed that the merits of a pictorial display outweigh the shortcomings and that the deficiencies discussed point the way toward developing a useful display without relying only on additional aircraft type instrumentation. For example, better knowledge regarding glide slope errors can be provided by adding a mark displaced 3° below the horizon. When the aircraft is flown so that this indicator overlays a point on the runway, the pilot will know that he is on a 3° glide slope to that point. It is believed that this and other concepts in visual displays may remove the deficiencies noted, and research along these lines is in progress.

ACKNOWLEDGMENT

The writers appreciated the cooperation of the four commercial aviation pilots who participated in this study.

These pilots, who were selected to represent a broad spectrum of pilots competent in instrument flying, devoted their time and efforts principally from an interest in and a desire to contribute to the development of advanced airborne approach and landing aids.

APPENDIX A

AIRCRAFT SIMULATION

The following linear differential equations were programmed on a digital computer to simulate the dynamics of a Navion, a single-engine, four-place light aircraft.

Airframe Dynamics

$$\begin{bmatrix} \dot{u} \\ \dot{w} \\ \dot{q} \end{bmatrix} = \begin{bmatrix} X_u & X_w & -g/s \\ Z_u & Z_w & U_o \\ O & M_w & M_q \end{bmatrix} \cdot \begin{bmatrix} u \\ w \\ q \end{bmatrix} + \begin{bmatrix} O \\ Z_{\delta_e} \\ M_{\delta_e} \end{bmatrix} \delta_e$$

$$\dot{p} = L_p \cdot p - L_{\delta_\alpha} \cdot \delta_\alpha$$

The following values of the stability derivatives were obtained from reference 11.

Z_{δ_e}	= -28.17	ft $(sec)^{-2}$	M_w	= -0.04997	$(sec\ ft)^{-1}$
M_{δ_e}	= -11.1892	$(sec)^{-2}$	M_q	= -2.0767	$(sec)^{-1}$
X_w	= 0.03607	$(sec)^{-1}$	U_o	= 176.0	ft $(sec)^{-1}$
X_u	= -0.0451	$(sec)^{-1}$	L_p	= -8.402	$(sec)^{-1}$
Z_w	= -2.0244	$(sec)^{-1}$	L_{δ_α}	= 23.984	$(sec)^{-2}$
Z_u	= -0.3697	$(sec)^{-1}$			

The equations of motion were solved every 0.05 seconds, with the necessary integrals being approximated by equivalent sums. For example:

$$u_{new} = \dot{u} * 0.05 + u_{old}$$

The following approximation for yaw rate (r) resulted in coordinated turns for small bank angles.

$$r = p \cdot g/(U_o \cdot s)$$

Control forces. - A MSI Model 438 2-axis side-arm force controller with a flexible fiber control stick was used. The fiber stick had a 16 pound per inch restoring force. The stick gains were:

longitudinal: 25.4 lb./rad. of δ_e

lateral: 15.6 lb./rad. of δ_a

Coordinate Transformations

Approximate Euler Transformations. -

$$\dot{\Theta} = q \cdot \cos(\phi) - r \cdot \sin(\phi)$$

$$\dot{\psi} = q \cdot \sin(\phi) + r \cdot \cos(\phi)$$

$$\dot{\phi} = p$$

Ground coordinate approximations. -

$$\dot{X} = U_o \cdot \cos(\Theta) \cdot \cos(\psi)$$

$$\dot{Y} = U_o \cdot \sin(\psi) + \text{crosswind velocity}$$

$$\dot{Z} = U_o \cdot \sin(\Theta) + (w + 0.05236 \cdot U_o) \cdot \cos(\psi) + \text{draft velocity}$$

APPENDIX B

INSTRUCTIONS TO PILOT-SUBJECTS

Introduction. - You are about to participate in an experiment aimed at evaluating a new kind of IFR approach and landing aid. The basic feature of this proposed system is an airborne microwave receiver that is capable of sensing the location of certain kinds of markers placed on the airport terrain and displaying a perspective view of these markers on a cathode ray tube mounted in the aircraft instrument panel. At its very best, this system would present a picture that could be similar to that of a runway viewed from an aircraft on a dark night.

However, there are certain predicted shortcomings of this yet to be built system; namely, the display resolution in the vertical and the horizontal directions will be less than perfect, and there will be a time period required to process the received data and make the picture. The purpose of this experiment, using experienced instrument pilots from the general aviation and commercial aviation groups, is to determine how these display variables will affect performance in the approach and landing task. These findings will be used to establish specifications for a prototype system, to estimate the potential applications of the system and to point out inherent weaknesses of the system that may be improved by relocating and shielding the airport markers.

Potential applications of this system are:

1. A low cost IFR approach system that can be installed at any airport.
2. A zero visibility (Category III) manual approach and landing system.
3. An independently derived visual back-up for a fully automated landing system.

Please read your instructions carefully and be sure that the NASA project engineer answers to your satisfaction any questions that you may have regarding your participation in this experiment. You will be given a copy of the NASA report on this experiment when it is ready for publication.

Experiment I. -

1. You are to fly a simulation of a Navion (modified for two-control) down an imaginary flight path and land using only the display provided. You will be positioned initially on a three degree glide-slope that intersects the runway 1000 feet down the runway from the threshold. There will be random winds which will affect your altitude and drift over the first portion of your approach and then they will cease suddenly. At all times you are to make every effort to stay lined up with the runway centerline and on a three degree glide-slope. Though the engine power is set for the approach, you are to flare and try to make a successful landing at a point 1200 feet down the runway. (Each pilot was provided with a copy of figure 2 of this report - a diagram of the runway and approach profile.)

2. You are instructed to land the simulated aircraft on each run; however, if you reach a point where you normally would abort the landing, press the switch to the left of the display but continue to land the "aircraft." After landing, indicate on the rating sheet why you would have aborted.

3. You are to use the rating sheet provided to rate the quality of each display after each set of runs with that display for:

a. The approach up to the first lead-in light.

b. From the first lead-in light up to the landing or the point at which you would have aborted the landing if you were able to.

(The computer operator will inform you when to rate the display for a set of runs.)

4. Your performance on each approach and landing will be scored by measuring vertical and lateral errors at points 5000 feet, 1000 feet and 0 feet from the runway threshold. Also, scores will be obtained for sink rates, lateral errors from the runway centerline and distances from the 1200 foot spot landing point at touchdown.

5. Please study the rating criteria sheet before the test runs. A copy is provided in the cab for your reference during the runs.

Experiment II. - For these flights you are to approach and land as before; however, you will note that the horizon will be missing from the display.

It is emphasized that at all times you are to make every effort to stay lined up with the runway centerline and on a three degree glide-slope.

The instructions for the rating sheet and the use of the abort button are the same as before.

Experiment III. - In this experiment you will be placed a fixed distance from the runway with a considerable error in either height or lateral position, but not both. You are to move the controller in the proper direction to shift your aircraft up or down, or right or left, as the case may be, to position yourself on the correct flight path as established in the prior experiments. Then move your aircraft further in the same direction until you feel that you have just gone past the proper flight path, and if you were flying by use of the display, you would make a correction back in the opposite direction. At this time press the switch at the left of the display to indicate your decision that you are just off of the proper flight path in the opposite direction from the position at which you were started.

You will only be able to move your aircraft in one direction from the initial large error in position, that is, you cannot back up by reversing the controller.

APPENDIX C

DEBRIEFING

At the culmination of the three experiments, each pilot was individually asked a number of questions and the responses recorded. The following paragraphs summarize these interviews.

Question 1. - How did the workload of the approach task with the display that had the horizon on it compare with an ILS approach?

One of the pilots said that the approach with the pictorial display was harder (". . .I'm used to an ILS," though, ". . .I think that someone just learning may find the display easier."); another stated that the task was the same in workload up to the lead-in lights, then easier; and the remaining two pilots said that the pictorial display was easier to use.

Question 2a. - What are your normal glideslope errors in making an ILS approach?

The pilots stated that they were able to keep the glideslope indicator within 25% to 50% of the full scale deflection, with two of them stating that they stayed on the high side of the established glide path.

Question 2b. - What are your normal localizer errors in making an ILS approach?

All of the pilots stated that they were able to maintain the localizer needle within 25% of full scale.

Question 3. - Did you develop any special strategy in using the display?

Three of the pilots reported that they intentionally flew below what they thought was the proper glideslope, though their reasons for doing this differed (one did it to get a better picture, one felt that he had to dive a little to touchdown at the 1200 foot distance, and the other was overcoming what he believed to be a natural tendency to be too high due to the type of display). One of the pilots reported that he was able to anticipate the abatement of the crosswinds and to initiate a recovery from the crab attitude by noting the distance to the lead-in lights without waiting for a lateral displacement to occur.

Question 4a. - How would you like the display improved in configuration?

All of the pilots, in one way or another, indicated that they wanted the display to contain better information from which the distance that the aircraft had traveled down the runway could be determined. Two of the pilots wanted horizontal lines (roll bars)

at the threshold and at the 1,000 foot mark. One of the pilots wanted a centerline on the last two-thirds of the runway. One of the pilots wanted an abort mark placed where insufficient runway was remaining.

Question 4b. - How would you like the display improved to give more information?

All of the pilots requested additional instrumentation. Three wanted the instruments off to one side within the field of view. Instruments requested were as follows:

Instrument	Number of pilots requesting
Rate of climb	3
Altimeter	3
Airspeed	2
Heading	1
ILS indicator	1
Turn and bank	1
Power	1

Question 5a. - What was the most disturbing aspect of the display during the approach?

Each of the pilots responded differently. One said that the sink rate could not be determined and too much correction was required at the lead-in lights (recovery from crab in crosswind); another said that with the slow update interval, he tended to overcontrol and lacked confidence; another said that with the degraded resolution and lack of definition, he had trouble determining the center of the runway; and finally, the remaining pilot said that when the display had discrete jumps, it was hard to control heading and that a compass was needed.

Question 5b. - What was the most disturbing aspect of the display during the landing?

The pilots again varied in their responses. One was disturbed by the lack of centerline information when close to the runway; another said that the effect of a slow update interval, which caused tension and overcontrolling, was even worse during the landing than during the approach; another was disturbed by the sensitivity of the controller, particularly in roll; and the remaining pilot was disturbed by his tendency to land long and inability to determine altitude from the display.

Question 6. - What did you think of the simulation of a "two-control Navion?"

Most of the pilots thought that the simulation was adequate, though two complained about the control being too sensitive, particularly in roll. However, one thought that the simulation was heavier on the controls than a Cessna 182.

Question 7. - Did you have enough practice to do the approach and landing task?

All of the pilots stated that enough or more than enough practice was given.

Question 8. - Do you think that your performance would improve much with more practice?

All pilots answered in the affirmative.

Question 9. - Did you have difficulty in seeing the picture of the runway as being a real runway?

None of the pilots said that he had any difficulty in seeing the display as a view of a runway on a dark night when the resolution and update interval of the display was good; however, this was not true when the picture was degraded.

Question 10. - Was the task with the display more or less fatiguing than a similar ILS approach?

Responses were similar to those of question 1, concerning workload.

Question 11. - Were your instructions clear?

All of the pilots said that their instructions were clear prior to the data runs, though one believed that a more basic explanation of how to use the display would have been appropriate at the beginning.

Question 12. - Would you desire to participate in future experiments like this one?

All answered in the affirmative.

REFERENCES

1. Robinson, G. and Johnson, N.: Subsystem Requirements for an Airborne Laboratory to Study Zero-Zero Landing Systems. NATO AGARD Report No. 488, 1964.

2. Young, D. and Suzansky, J.: Research Study of An Aircraft-Contained Radar Zero-Zero Landing System, Vol. 1. NASA CR 73184, 1967.

3. Ketchel, J. and Jenney, L.: Electronic and Optically Generated Aircraft Displays. JANAIR Report No. 68505, May 1968.

4. Cooper, G. and Harper, Jr., R.: The Use of Pilot Rating in the Evaluation of Aircraft Handling Qualities. NASA TN D-5153, 1969.

5. McRuer, D., Ashkenas, I. and Pass, H.: Analysis of Multiloop Vehicular Control Systems. Tech. Documentary Report No. ASK-TDR 62-1014, March 1964.

6. Kraft, C. and Elworth, C.: How High is Up. Interceptor, pp. 6-14, October 1968.

7. Anon: Automatic Landing Systems. FAA Advisory Circular No. 20-57. Effective January 29, 1968.

8. Hyman, A. and Gold, T.: Dynamic Visual Cues in Flying. (In Current Developments in Optics and Vision, Benson, W., Editor) Armed Forces - NCR Committee on Vision. Washington, D.C., 1968, pp. 3-16.

9. Anon: Criteria for Approval of Category II Landing Weather Minima. FAA Advisory Circular No. 120-20, Effective June 6, 1966.

10. Kibort, B. and Drinkwater III, F.: A Flight Study of Manual Blind Landing Performance Using Closed Circuit Television Displays. NASA TN D-2252, 1964.

11. Teper, G.: Aircraft Stability and Control Data. NASA CR 96008, STI TR 176-1, April 1969.

Table 1. — Experiment I height and lateral position statistics at a distance of 1,000 feet from the threshold. 26 flights per statistic (a positive lateral error is in the direction blown by the crosswind).

Pilot	Draft or Crosswind (ft/s)	Average Height (ft)	Average Error (ft)	σ (ft)
		Height results		
	-3	74	-31**	26.
1	0	104	- 0	40
	+3	127	+23**	22
	-3	68	-37**	11
2	0	97	- 8**	8
	+3	118	+13*	25##
	-3	43	-62**	12
3	0	95	-10**	14
	+3	167	+62**	21#
	-3	23	-81**	10
4	0	89	-16**	12
	+3	172	+87**	17
Pitch only	-3	20	-85	--
auto-pilot	0	105	0	--
	+3	190	+85	--
		Lateral position results		
1	0	--	+ 1	6
	9	--	- 2	22##
2	0	--	- 5**	10
	9	--	+21**	14#
3	0	--	- 4*	8
	9	--	+ 4	22##
4	0	--	- 1	14
	9	--	-21	64##
Heading only	0	--	0	--
auto-pilot	9	--	+256	--

*Significantly different from zero at .05 level of confidence.
**Significantly different from zero at .01 level of confidence.
#Sigma significantly larger than sigma for no wind flights at .05 level.
##Sigma significantly larger than sigma for no wind flights at .01 level.

Table 2. — Qualitative definition of touchdown performance. The rating for each landing was the worst rating of any of the three variables — L, D or S.

Touchdown performance description	Lateral displacement from centerline L, ft	Distance down runway D, ft	Sink rate S, ft/s
Successful	$\|L\| < 36$	$600 < D < 2300$	$S < 5$
Marginal	$36 < \|L\| < 72$	$0 < D < 600$ or $2300 < D < 3400$	$5 < S < 8$
Unsuccessful	$72 < \|L\|$	$D < 0$ or $3400 < D$	$8 < S$

Table 3. — Comparison of pilot ratings (Cooper-Harper scale) with and without horizon on display.

Display condition			Flight phase	Average pilot rating		Difference in average rating (a - b)
HR	VR	UI		Without horizon (a)	With horizon (b)	
.1	.1	.05	approach	4.9	4.0	0.9
			touchdown	6.0	4.5	1.5
.4	.8	.3	approach	8.1	5.3	2.8
			touchdown	9.8	5.8	4.0
.1	.1	1.0	approach	8.3	6.5	1.8
			touchdown	9.3	7.3	2.0
.4	.8	1.0	approach	7.7	6.5	1.2
			touchdown	9.3	6.0	3.3

Table 4. — Results of vertical position judgments for displays (HR = .1, VR = .1) and (HR = .4, VR = .8) combined. Statistics were derived from 8 judgments per pilot.

Pilot	Distance from threshold							
	1,000 feet				5,000 feet			
	Ave. ht.		Sigma ht.		Ave. ht.		Sigma ht.	
	(deg)	(ft)	(deg)	(ft)	(deg)	(ft)	(deg)	(ft)
			Judgments from above					
1	3.45	120	.36	12	3.59	376	.40	42
2	3.02	105	.47	16	3.70	387	.32	34
3	3.60	126	.34	12	3.70	387	.29	30
4	3.05	106	.50	17	3.58	375	.30	31
			Judgments from below					
1	3.07	107	.54	19	3.19	334	.30	31
2	3.06	107	.14	5	3.40	356	.31	32
3	4.70	164	.32	11	5.09	533	.52	54
4	3.59	125	.54	19	4.18	438	.58	61

FIGURE LEGENDS

Figure 1. Simulation equipment.

Figure 2. Runway and approach profile.

Figure 3. Digital display with the best resolution. Aircraft is 1,500 feet from the threshold, 22 feet right of centerline, 152 feet above fround level, pitched down 1°, banked 5° left and headed 1° left. Note: The altimeter and rate of climb instruments were present only during practice.

Figure 4. The 26-display conditions used in Experiment I. The various conditions are perturbations in one variable at a time about either of the two conditions: (HR = 0.1, VR = 0.1, UI = 0.05) or (HR = 0.4, VR = 0.8, UI = 0.30).

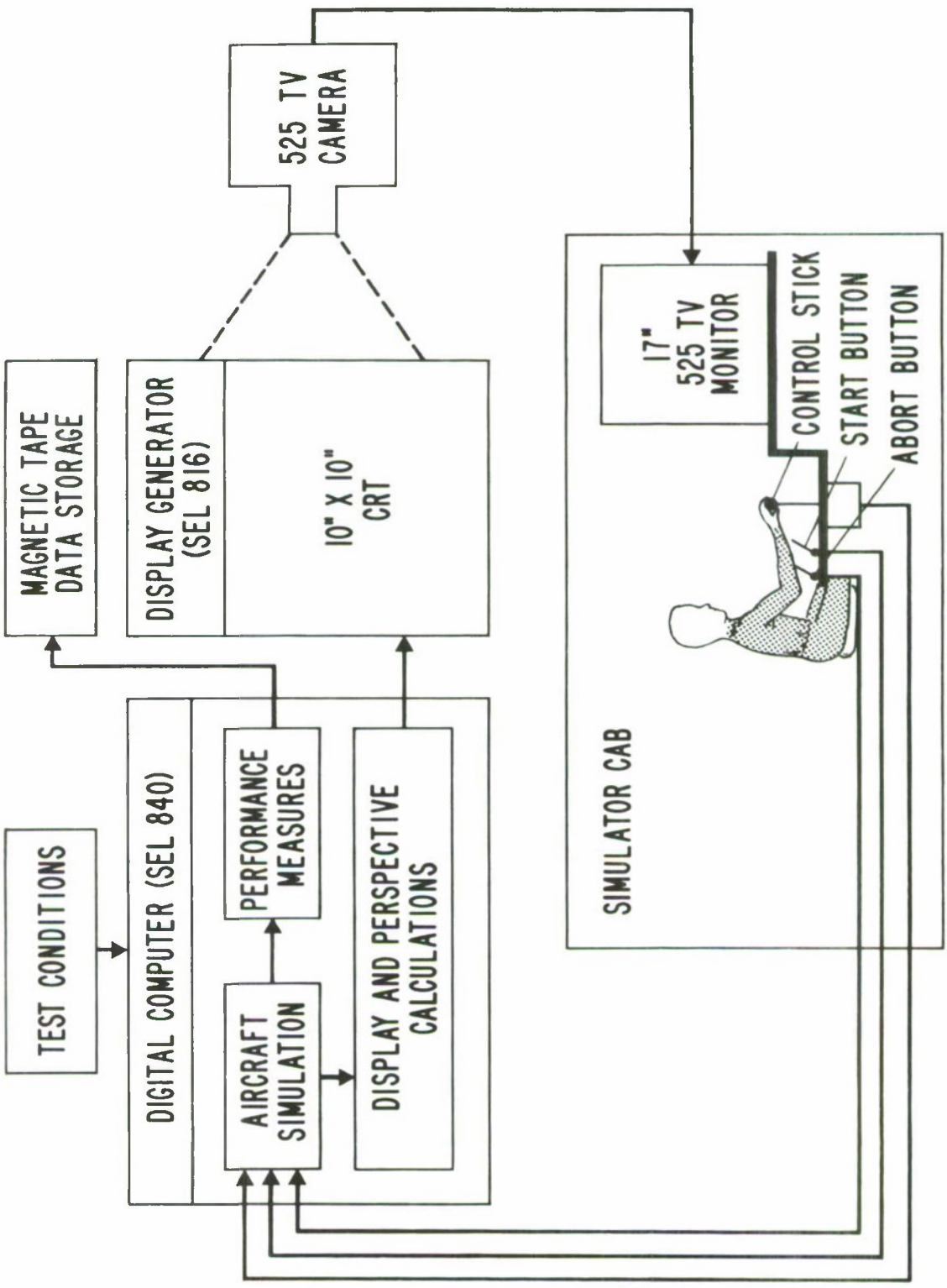
Figure 5. Pilot touchdown performance; all display conditions.

Figure 6. Touchdown performance by display condition for pilots 1, 2, and 3; 12 flights per condition.

Figure 7. Average Cooper–Harper ratings for each display condition; average range at each condition = 2.5 units.

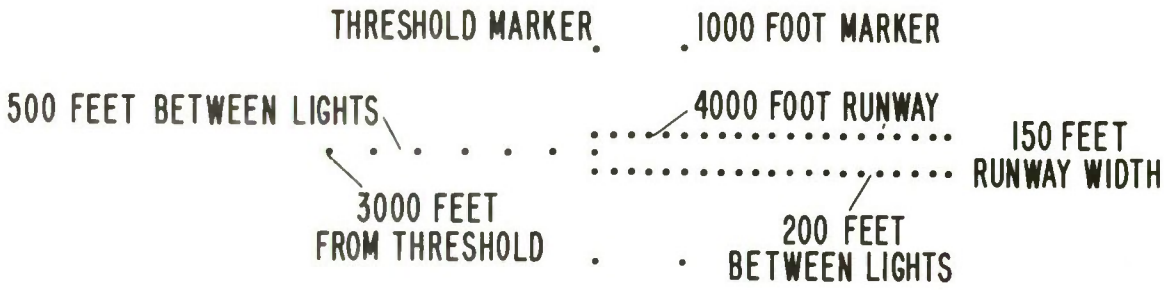
Figure 8. Touchdown performance with and without horizon on display.

Figure 9. Vertical position judgments for displays (HR = 0.1, VR = 0.1) and (HR = 0.4, VR = 0.8) combined. Each plot is mean $\pm 2\sigma$ for eight estimates per pilot of a 3° glide path. Arrows indicate direction of motion in making judgment.

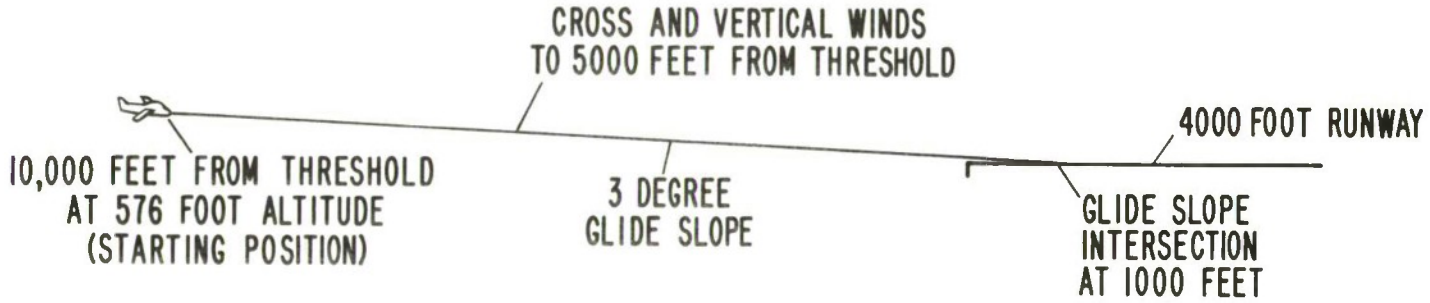


NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
AMES RESEARCH CENTER, MOFFETT FIELD, CALIFORNIA

Figure 1.



(a) Runway configuration



(b) Glide slope plan

Figure 2.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
AMES RESEARCH CENTER, MOFFETT FIELD, CALIFORNIA

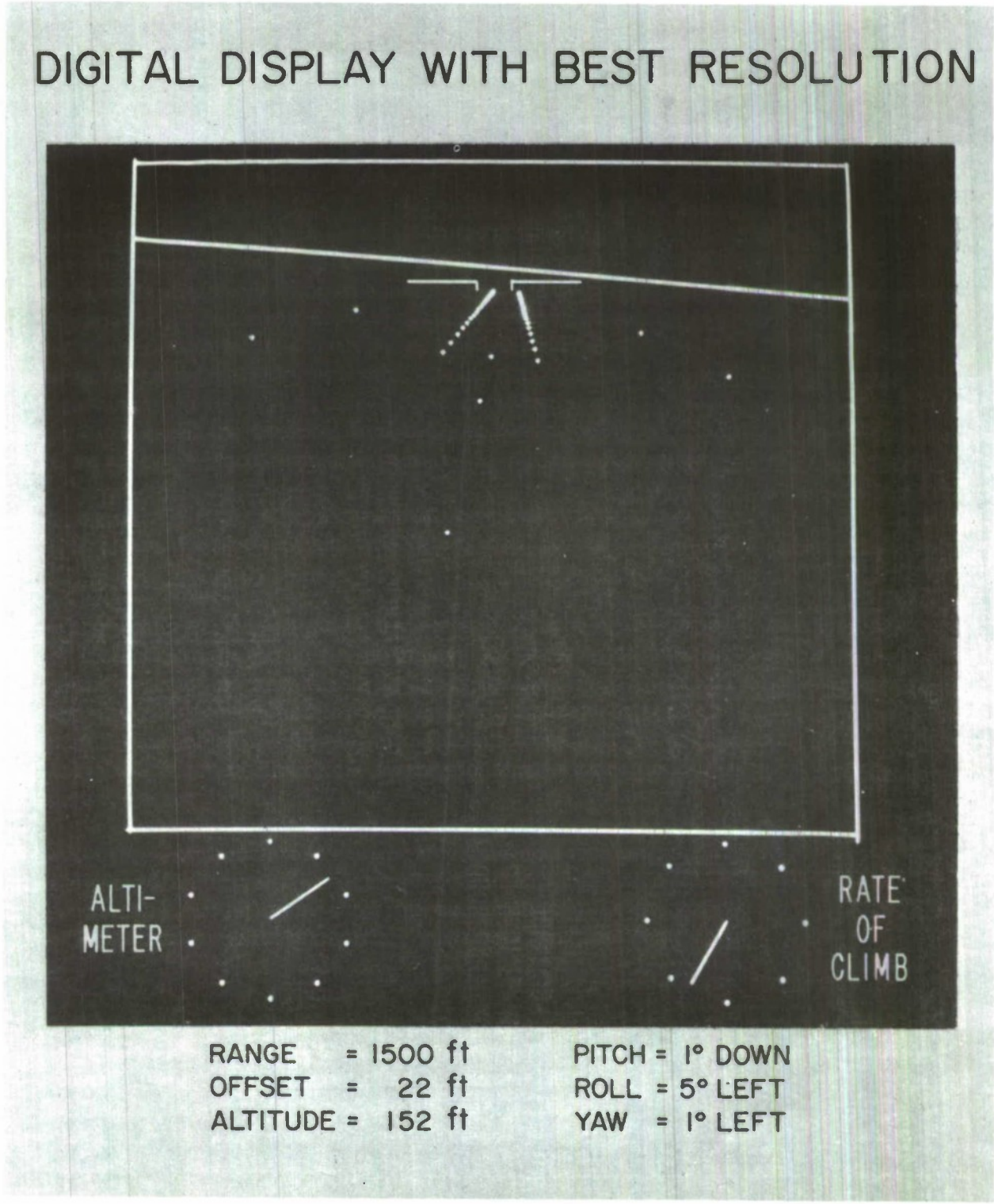


Figure 3.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
AMES RESEARCH CENTER, MOFFETT FIELD, CALIFORNIA

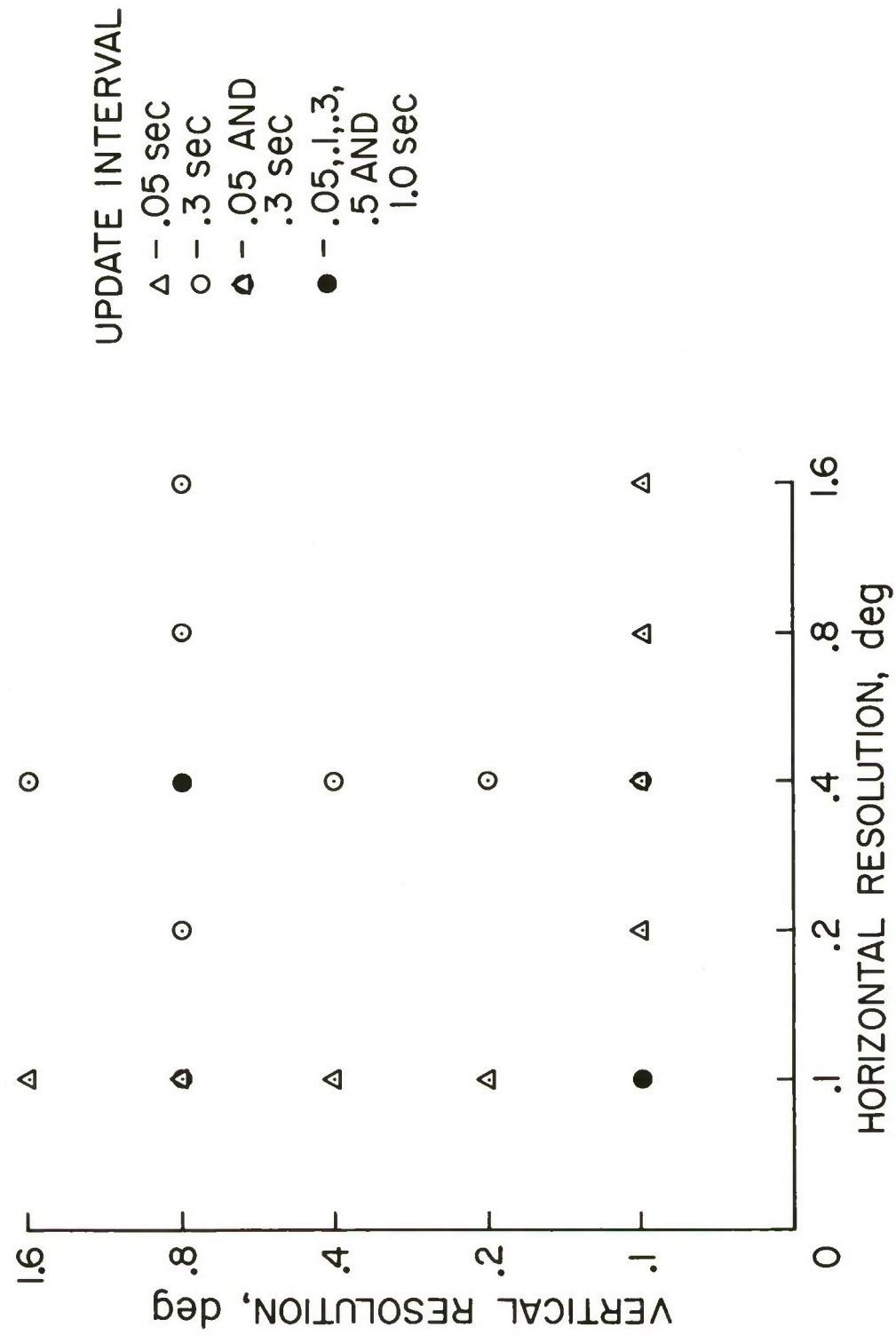
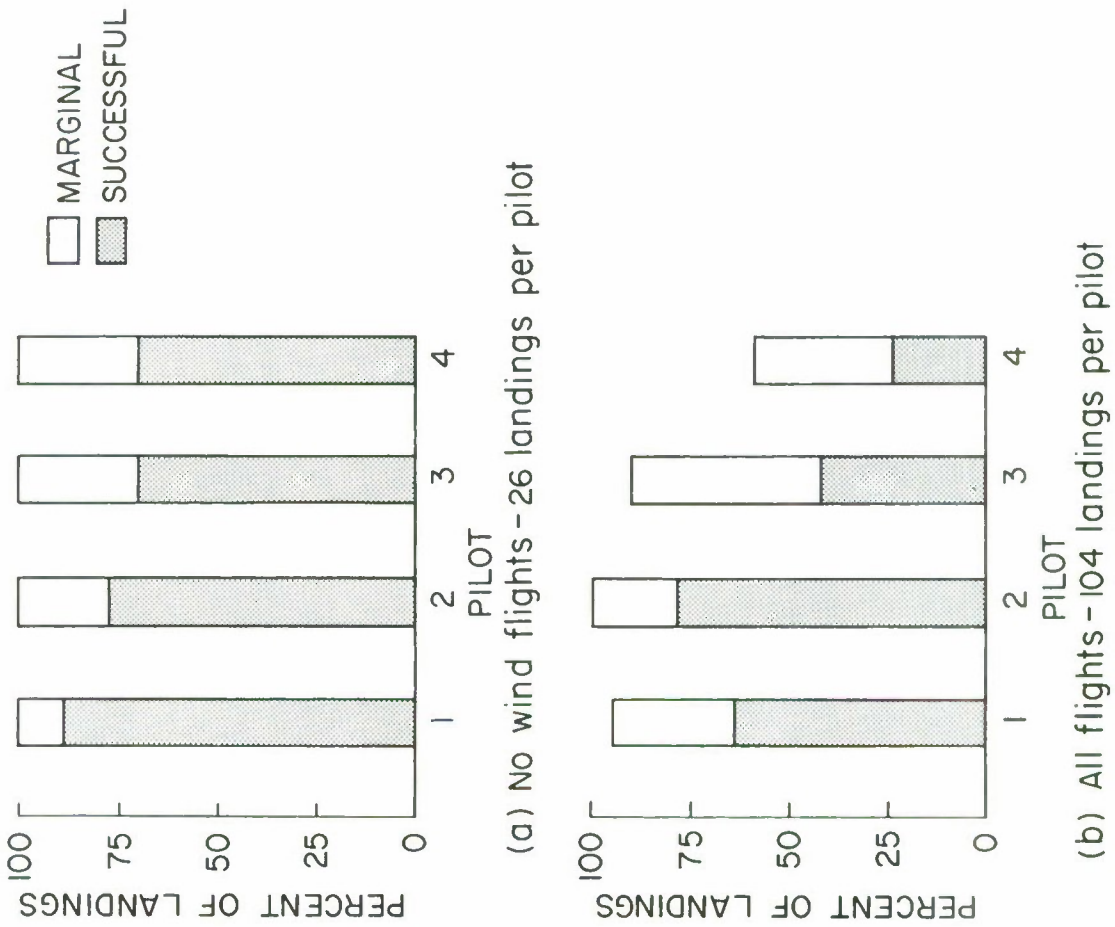

UPDATE INTERVAL
△ – .05 sec
○ – .3 sec
◐ – .05 AND .3 sec
● – .05,.1,.3, .5 AND 1.0 sec
VERTICAL RESOLUTION, deg
HORIZONTAL RESOLUTION, deg
0
.1
.2
.4
.8
1.6
.1
.2
.4
.8
1.6
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
AMES RESEARCH CENTER, MOFFETT FIELD, CALIFORNIA

Figure 4.



NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
AMES RESEARCH CENTER, MOFFETT FIELD, CALIFORNIA

Figure 5.

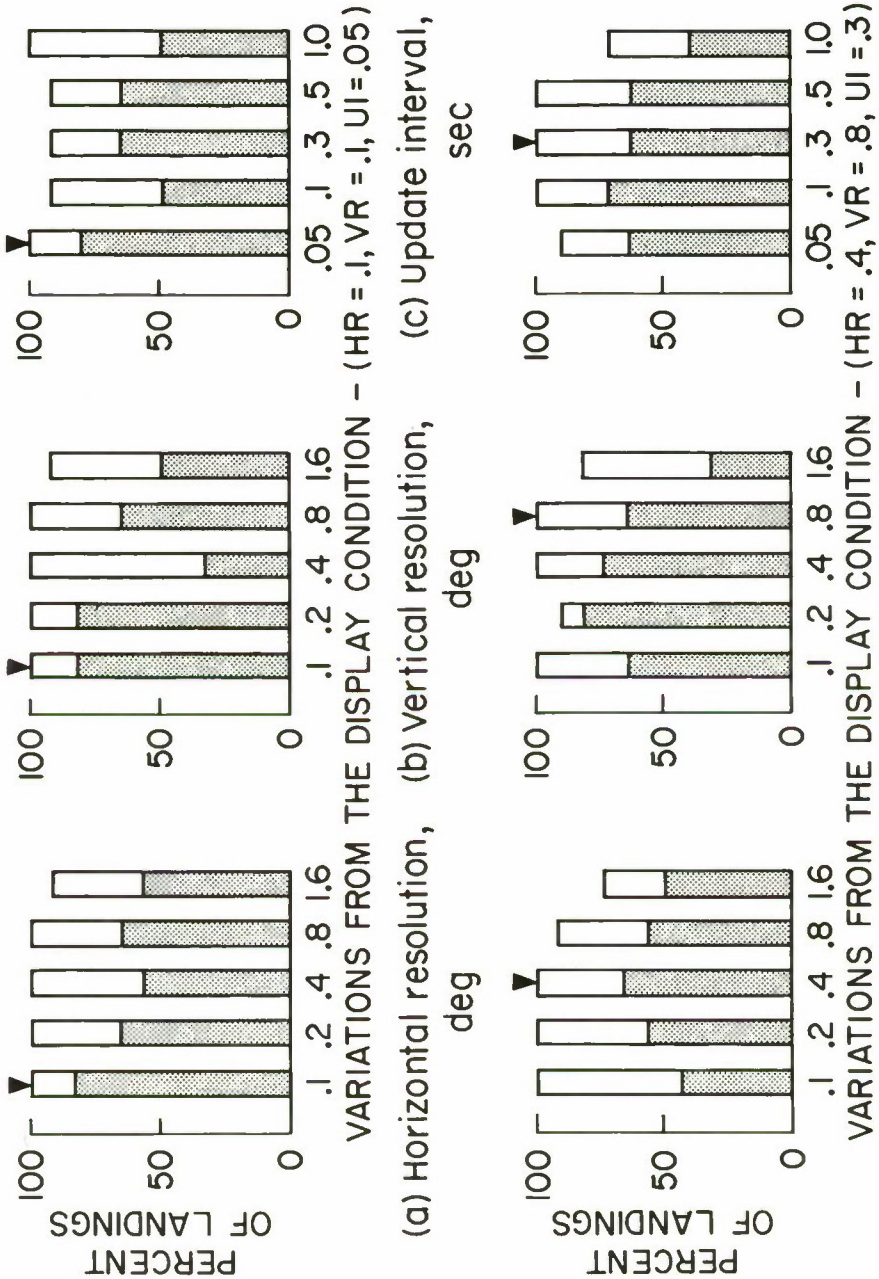


Figure 6.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
AMES RESEARCH CENTER, MOFFETT FIELD, CALIFORNIA

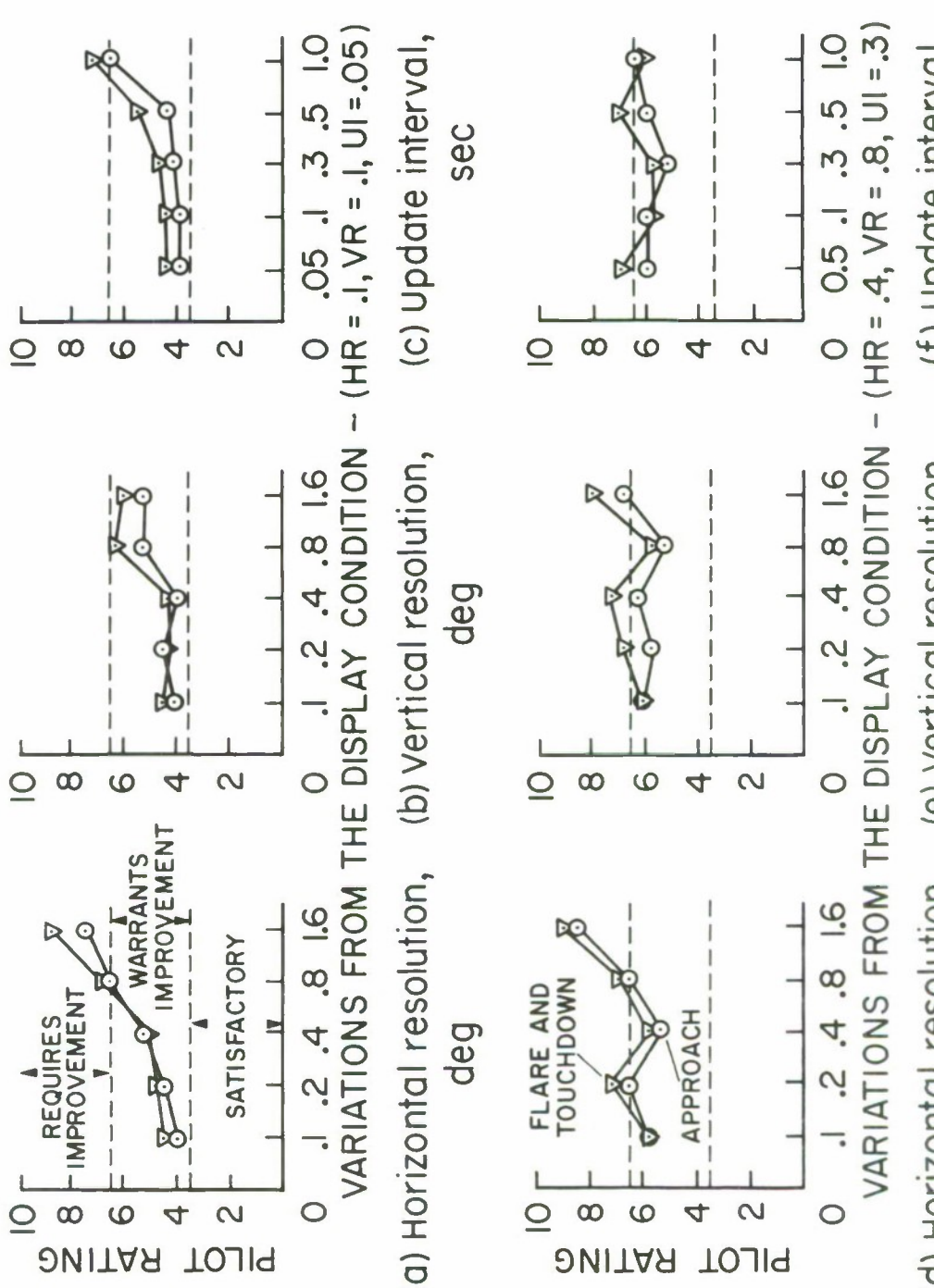
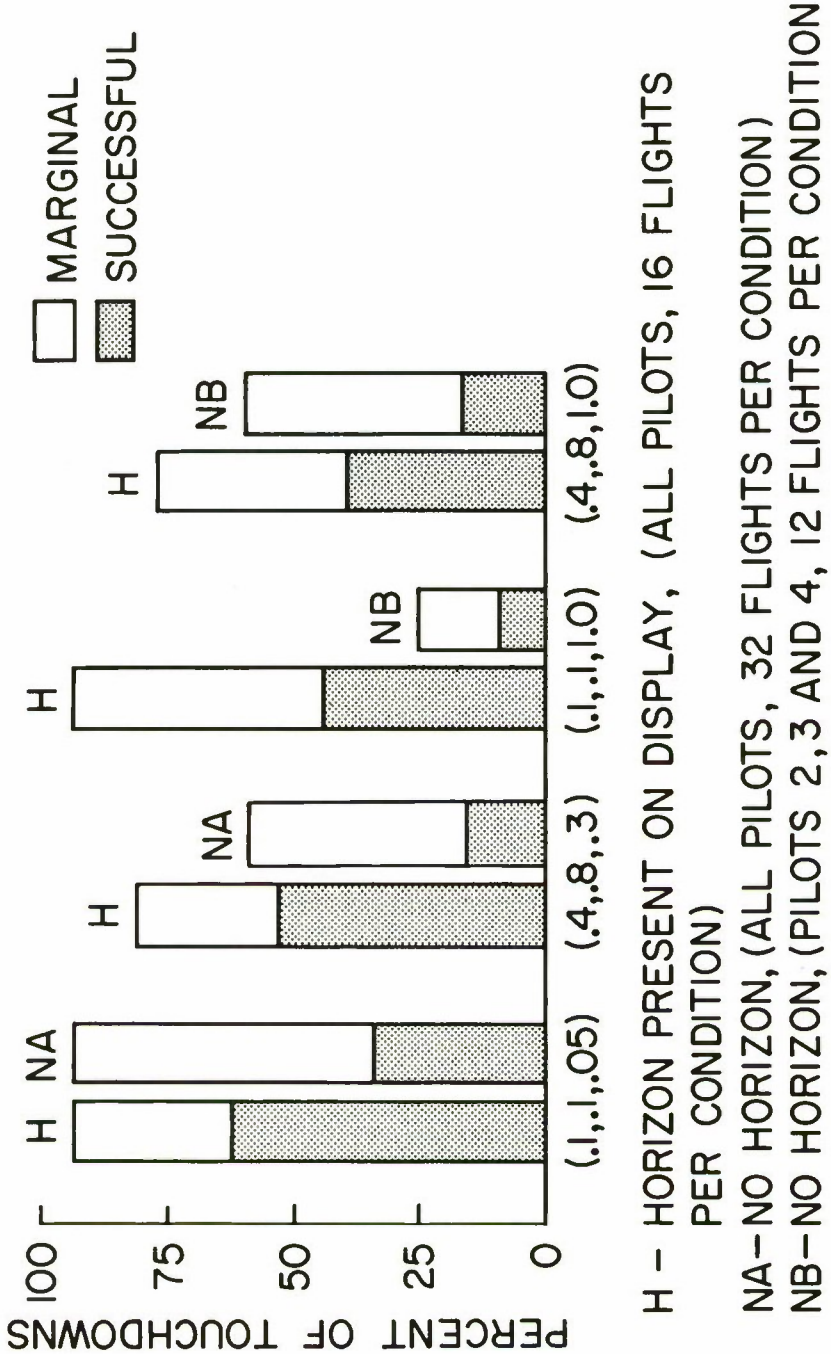


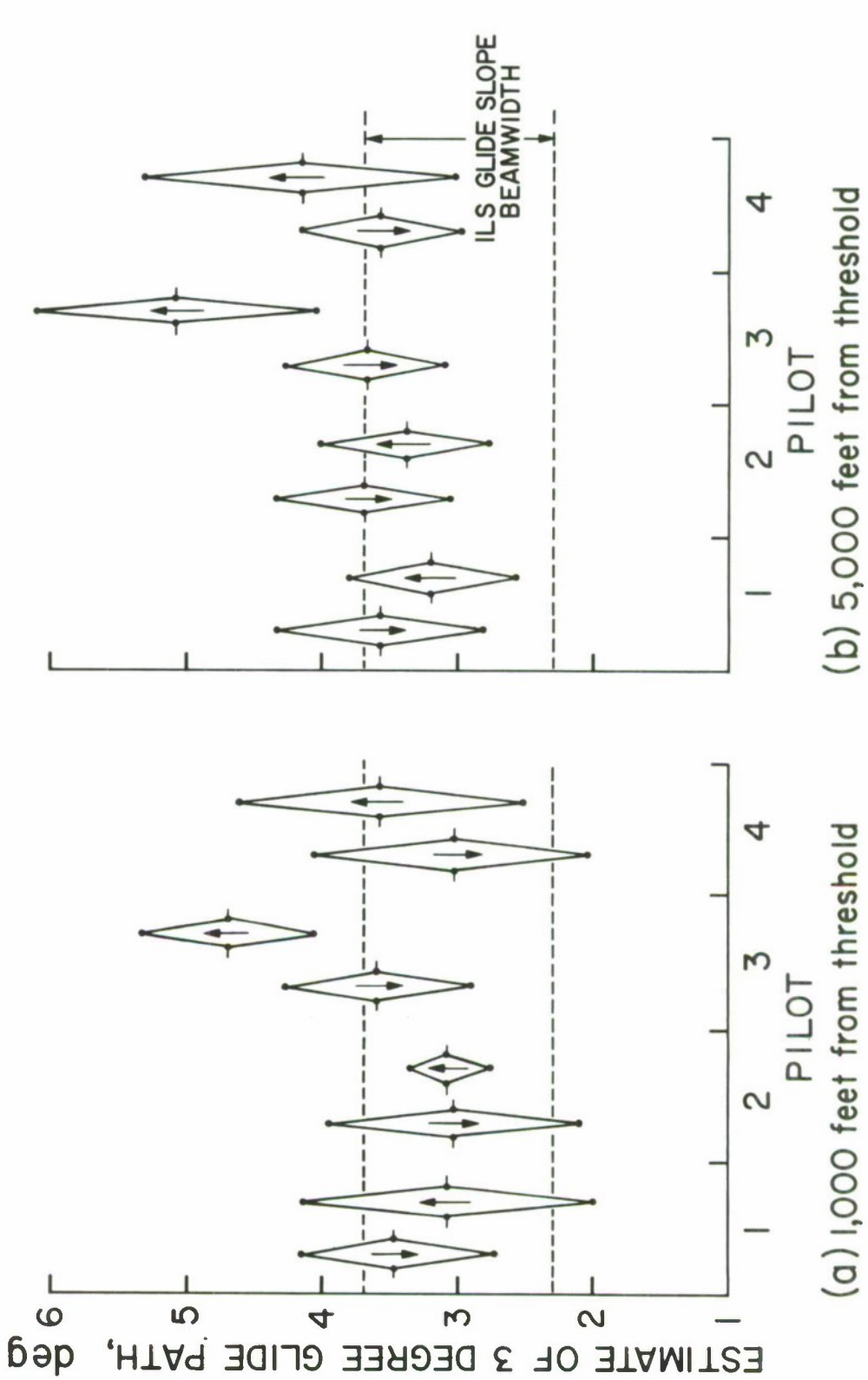
Figure 7.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
AMES RESEARCH CENTER, MOFFETT FIELD, CALIFORNIA



NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
AMES RESEARCH CENTER, MOFFETT FIELD, CALIFORNIA

Figure 8.



NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
AMES RESEARCH CENTER, MOFFETT FIELD, CALIFORNIA

Figure 9.

MEASUREMENT AND ANALYSIS OF PILOT SCANNING BEHAVIOR DURING SIMULATED INSTRUMENT APPROACHES

D. H. Weir and R. H. Klein
Systems Technology, Inc.
Hawthorne, California

ABSTRACT

Experimental measurements of pilot scanning and control response in a simulated instrument approach are reported. Airline pilot subjects flew ILS approaches in a six degree of freedom fixed-base DC-8 simulator at the NASA Ames Research Center. A conventional instrument panel and controls were used, with simulated vertical gust and glide slope beam bend forcing functions. Pilot eye fixations and scan traffic on the panel were measured using a recently developed eye-point-of-regard (EPR) system. Simultaneous recordings were made of displayed signals, pilot response, and vehicle motions. The EPR data were reduced for 31 approaches with a cross section of subjects to obtain dwell times, look rates, scan rates, and fractional scanning workload. Flight director (zero reader) approaches as well as standard localizer/glide slope (manual) approaches were made. The scanning results showed the attitude and glide slope/localizer instruments to be primary in a manual ILS approach, sharing 70 to 80 percent of the pilot's attention. The glide slope/localizer instrument required shorter dwell times with a fixed instrument sensitivity. Differences in dwell time between pilots occurred mainly on the attitude instrument. With the flight director, glide path deviation errors were reduced and the flight director instrument dominated pilot attention (about 80 percent). There were no apparent circulatory scanning patterns in any of the approaches. These EPR results were generally consistent with prior data where meaningful comparisons could be made.

INTRODUCTION

Further development and validation of the theory of manual control displays (Refs. 1 and 2) required measurement and analysis of simultaneous eye movement and pilot response data in flight control tasks under realistic instrument conditions. The objective of this research program* was to obtain such data for instrument approach tasks, and to reduce the eye point-of-regard (EPR) data to the scanning statistics needed to continue development of the theory. Data are now in hand for several airline pilots in more than a hundred simulated instrument approaches in a subsonic jet transport. The scanning statistics for a cross section of thirty-one 2 minute runs are described and analyzed in this paper. They are part of the data base for ongoing efforts to correlate EPR with pilot response and displayed motion variables.

BACKGROUND

Early research in this area was accomplished by Milton, Jones, and Fitts in an 8-year experiment. They used an eye camera to measure the instrument

*Accomplished in part under NASA Contract NAS2 -3746.

scanning patterns of pilots in a variety of actual IFR maneuvers (Refs. 4-8), but no records were made of the instrument readings or pilot responses. Stable statistical traffic patterns appeared in their results for various pilots and maneuvers. More recent research (e.g., Refs. 9-11) has been concerned mainly with statistical models of the scanning process, rather than with the establishment of connections with the causal factors of the displayed signals. Again, the displayed signals were either not recorded or not correlated against the scanning behavior.

The first effort to measure and correlate the visual sampling process and pilot control is reported in Ref. 12. It was aimed primarily at developing a queing theory for display scanning. Although detailed, these results are not suitable to validate the theory and methods of Refs. 1 and 2. Specific difficulties from our viewpoint include lack of a contemporary transport panel layout, no forcing functions (useful in measuring pilot response), and incomplete definition of the controlled element dynamics.

These past studies have used eye movement cameras, electro-oculographics, or corneal reflection techniques; which tend to be expensive, difficult to operate, and detrimental to the experimental environment. An eye-point-of-regard system recently developed by STI avoids many of these problems. It measures the horizontal and vertical movement of the eye with respect to the head by a corneal-scleral boundary contrast technique and the head movement relative to the panel reference by electromechanical means. These signals are combined in a special purpose computer to obtain overall eye point of regard. Figure 1 shows the device in use. This, plus proven experimental

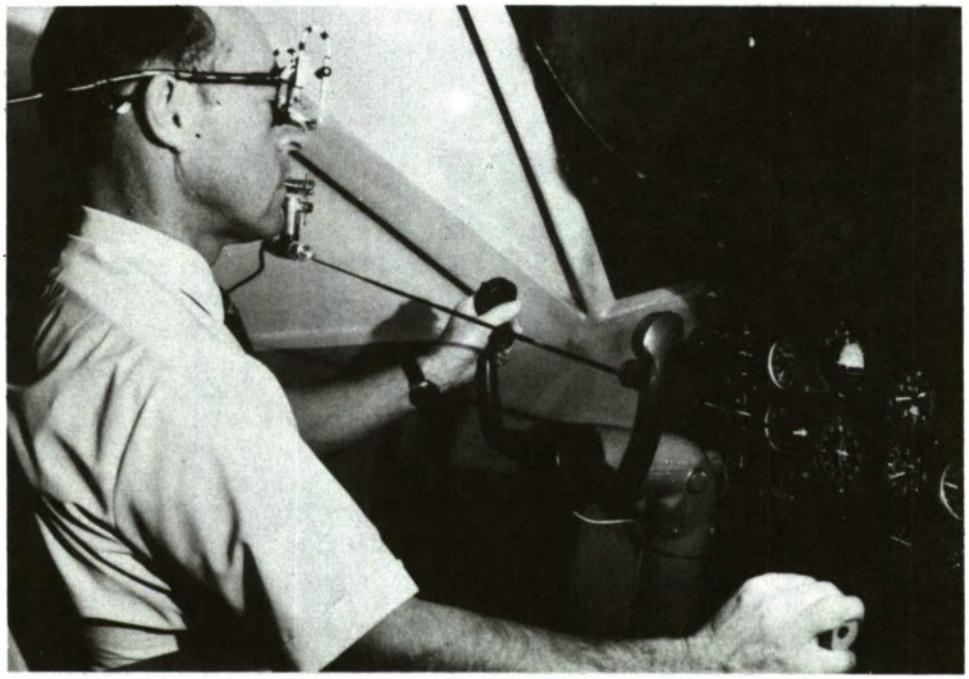


Figure 1. EPR System in Use

techniques for measuring pilot dynamic response in multiloop tasks (Refs. 13 and 14), and the availability of a high fidelity simulation facility at the NASA Ames Research Center, gave the tools to accomplish the behavioral measurement program.

EXPERIMENTAL SITUATIONS

This study used a NASA Ames fixed-base six degree of freedom simulator, configured as a subsonic jet transport. The task was to fly an Instrument Landing System (ILS) approach from the outer marker (30,000 ft from threshold) to the middle marker (2,500 ft from threshold) in the presence of vertical gusts and glide slope beam bends. Aircraft motions, displayed signals, pilot response, and eye point of regard were recorded. The system block diagram is shown in Fig. 2. Details of the experimental setup and procedures used are given in Ref. 3.

The experimental configurations are described in Table I. Configuration A was a pitch attitude only tracking task designed to provide single-loop response data on the present subjects for correlation with past data and models. For configurations B, C, and D the display included localizer and glide slope deviation, pitch and roll attitude, and peripheral instruments; but no flight director display. These tasks varied in their detail in order to explore effects of scanning and statistical stationarity. Configurations E and F employed all the displayes of C and D, respectively, plus a standard lateral and longitudinal flight director display superimposed on the artificial horizon instrument. The visual breakout runs involved transition from the panel display to the external visual field, inside the middle marker. The "fixed range" configurations had the instrument range varying sensitivities fixed at the outer marker.

The simulated vehicle was nominally a DC-8 jet transport in the landing approach configuration. The dynamics were defined by a linearized set of perturbation equations in six degrees of freedom. The simulator was stabilized with full flaps and gear down at 135 kts on the approach path at the outer marker at the start of the run. No changes in flaps, trim, or power setting were required during the run — although the pilot was free to make throttle corrections. The vehicle transfer functions and sample transient responses are given in Ref. 3.

The panel configuration was the standard "T," typical of current transport usage. The layout for the manual ILS configurations is shown in Fig. 3. The instrument needles have been deleted for clarity. The flight director bar appeared on instrument 2 for Configurations E and F. The longitudinal flight director mixed pitch attitude and altitude errors. The latter were computed from the angular glide slope deviation by multiplying by the range to the glide slope transmitter. This caused the forcing function amplitude (component due to the glide slope command) to decrease during Configuration F runs. The lateral director mixed roll angle, heading angle, and (angular)localizer deviation errors. A low gain "autopilot" was used in Configuration B to simulate human pilot control of the lateral axes. The panel instrument dynamics are part of the controlled element; and the attitude ball, glide slope bar, localizer bar, and pitch and roll director display frequency responses all looked like well damped second-order systems with break

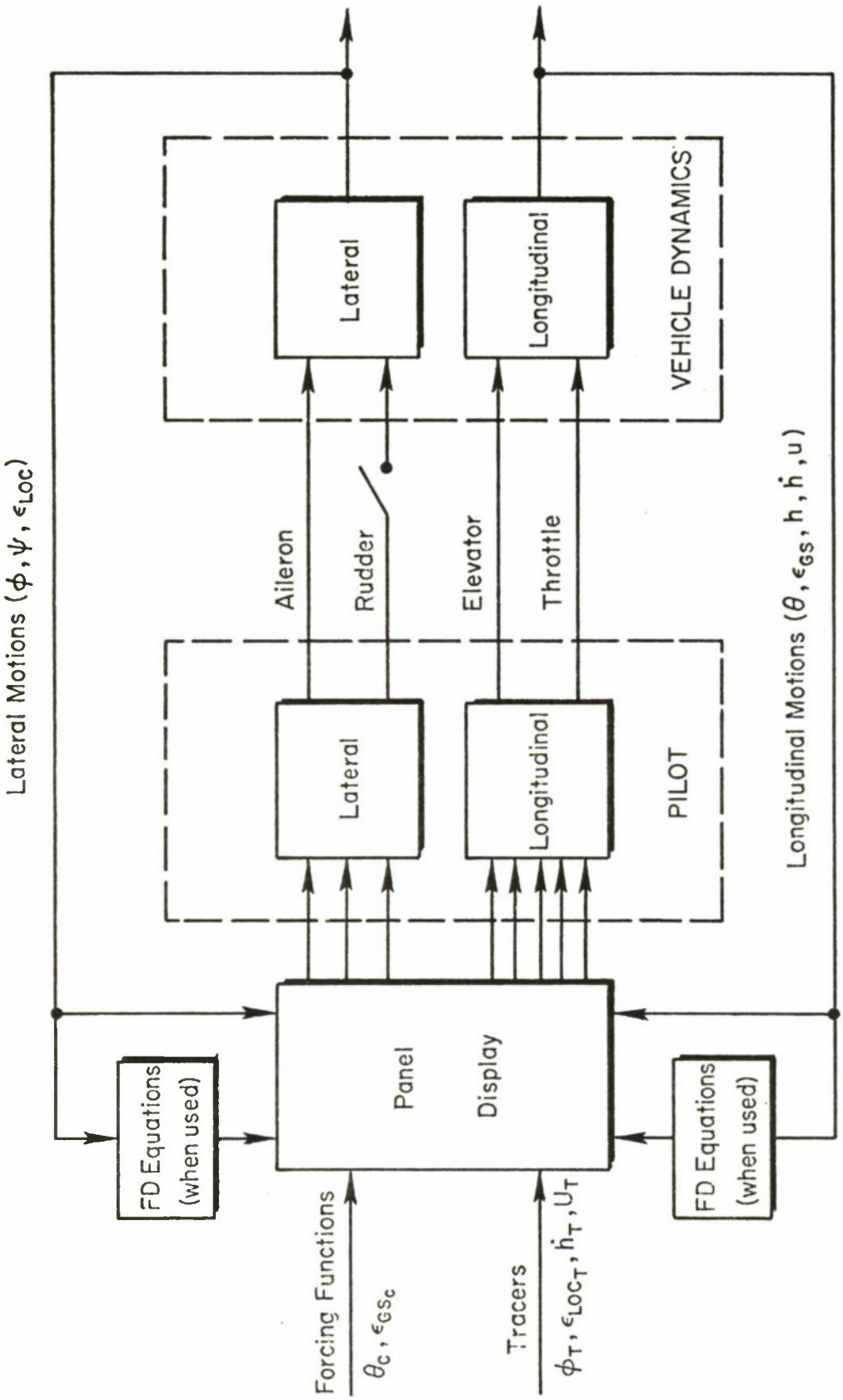


Figure 2. Block Diagram of Experimental Task

TABLE I

EXPERIMENTAL CONFIGURATIONS

CONFIGURATION	DESCRIPTION	PURPOSE	INSTRUCTIONS TO SUBJECTS
A (Pitch attitude regulation only)	Single axis tracking task with pitch attitude display and forcing function. Other instruments masked. Other axes controlled by autopilot. No flight director.	Tie in with single loop tracking data.	Simulates a portion of the approach task. Control pitch attitude only, and try to keep pitch error equal to zero. There is some turbulence. The lateral autopilot is ON.
B (Split axis manual ILS, fixed range)	Three degree of freedom longitudinal task. Forcing functions and tracers on. Lateral axes under autopilot control, but meters visible. No flight director.	Provide longitudinal scanning task, and basis for validating multiloop pilot response model.	Simulates a split-axis manual approach under Category II conditions. Control only the longitudinal motions. An autopilot is controlling the lateral motions. There is some turbulence. Try to keep the glide slope needle centered at all times.
C (Manual ILS, fixed range)	All axis approach task with forcing functions and tracers on. The glide slope deviation computer range was fixed at 30,000 ft from threshold. The altimeter and rate of climb meters appeared normal (varying range). No flight director.	Provides stationary all axis task. Reference case for comparison with split axis, range varying, and flight director cases.	Simulates a Category II manual ILS approach. There is some turbulence. Try to keep the glide slope and localizer needles centered at all times.
D (Manual ILS, varying range)	All axis approach task with forcing functions and tracers on. The range varied throughout the run. Glide slope deviation per unit altitude error increases with decreasing range. No flight director.	Provides nonstationary longitudinal task. Typical of "old fashioned" cross pointer ILS display.	
E (Flight Director, fixed range)	All axis approach task with forcing functions and tracers on. Flight director on, and driven by forcing function. Same as Configuration C plus flight director.	Provides equalized, integrated display and stationary all axis task. Reference flight director case. Typical of modern practice.	Simulates a Category II FD approach. There is some turbulence. Use the Director to follow the approach path, keeping the glide slope and localizer needles centered. Pitch commands must be obeyed immediately to avoid a standoff. The glide slope and localizer needles must be monitored.
F (Flight Director, varying range)	All axis approach task with forcing functions and tracers. Flight director on. Glide slope component of FD forcing function attenuated with range by flight director computer. Same as E, except range varying.	Provides equalized, integrated display but nonstationary forcing function.	
Visual Breakout	All axis approach task with Configurations D or F, except that the GPS system was used to provide an external visual field display at 200 ft above ground. Pilot looked up when told by copilot/experimenter. Breakout occurred some time after the 100 sec run.	Exploratory investigation of the feasibility of measuring both visual field and panel EPR. Used with only a few runs at the end of the experimental program.	Same as D or F, except that copilot calls out altitude and says "runway in sight" when visual acquisition occurs.

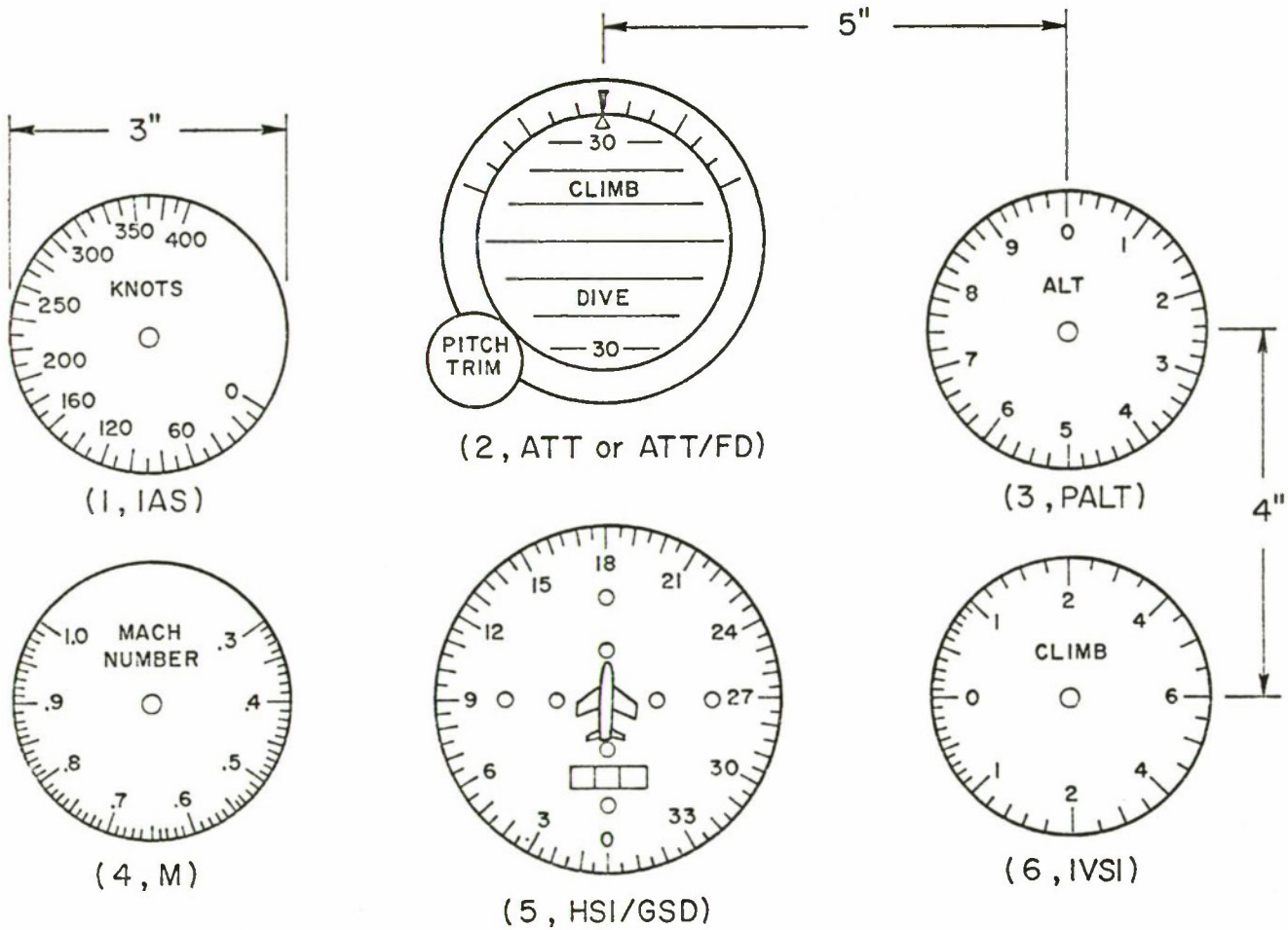


Figure 3. Laycut of Basic Flight Instruments

frequencies in the region 1 to 1.5 Hz. The peripheral instruments were more responsive. The column and wheel operated a hydraulic feel system. The DC-8 pilots termed it a reasonable facsimile of that aircraft, the Boeing 707 pilots felt it was somewhat light and sensitive, and the Convair 990 pilot thought it sluggish and insensitive.

Two independent longitudinal forcing functions, a pitch attitude command, θ_c, and a glide slope deviation command, ϵ_{GS_c}, were used in the experiments. This permits multiloop describing functions to be computed from the data. "Tracers," consisting of one or two low amplitude sine waves, were added to some other displayed signals.

The pitch attitude forcing function simulated a 5 ft/sec rms vertical gust input. It was a stationary random-appearing signal composed of a sum of sine waves. It had a bandwidth of 0.8 rad/sec and an rms amplitude of 1.2 deg, as shown in Fig. 4. Lower amplitude higher frequency components comprised a

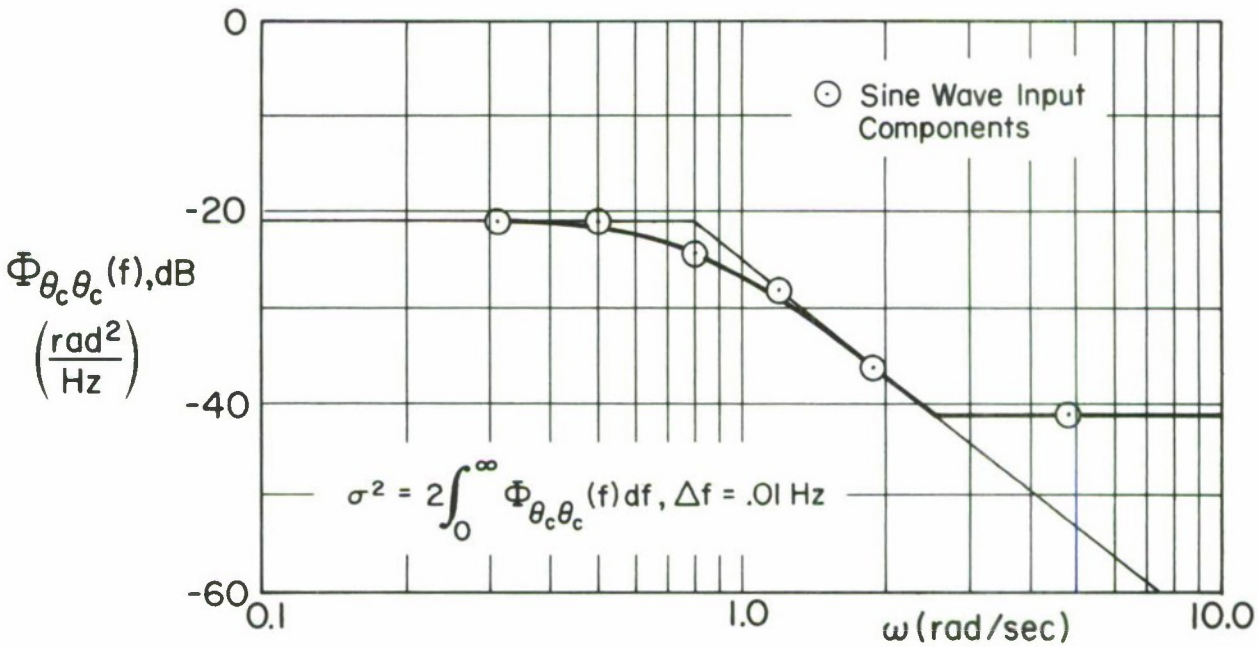


Figure 4. Power Spectral Density of Pitch Angle Forcing Function

"shelf" to facilitate response measurements in the expected region of the θ-loop pilot crossover frequency. The glide slope command forcing function simulated low frequency beam bends. It was a random-appearing sum of sine waves with an effective bandwidth of about 0.3 rad/sec and a mid-frequency low amplitude shelf, as shown in Fig. 5. It had an rms amplitude of 0.04 deg path angle or about 0.2 dots of needle deflection. This input magnitude was around the upper limit of acceptability for Category II beam bends, and occasionally exceeded it. The two command inputs were independent, containing different component sine waves which were "interleaved," as shown in Figs. 4 and 5.

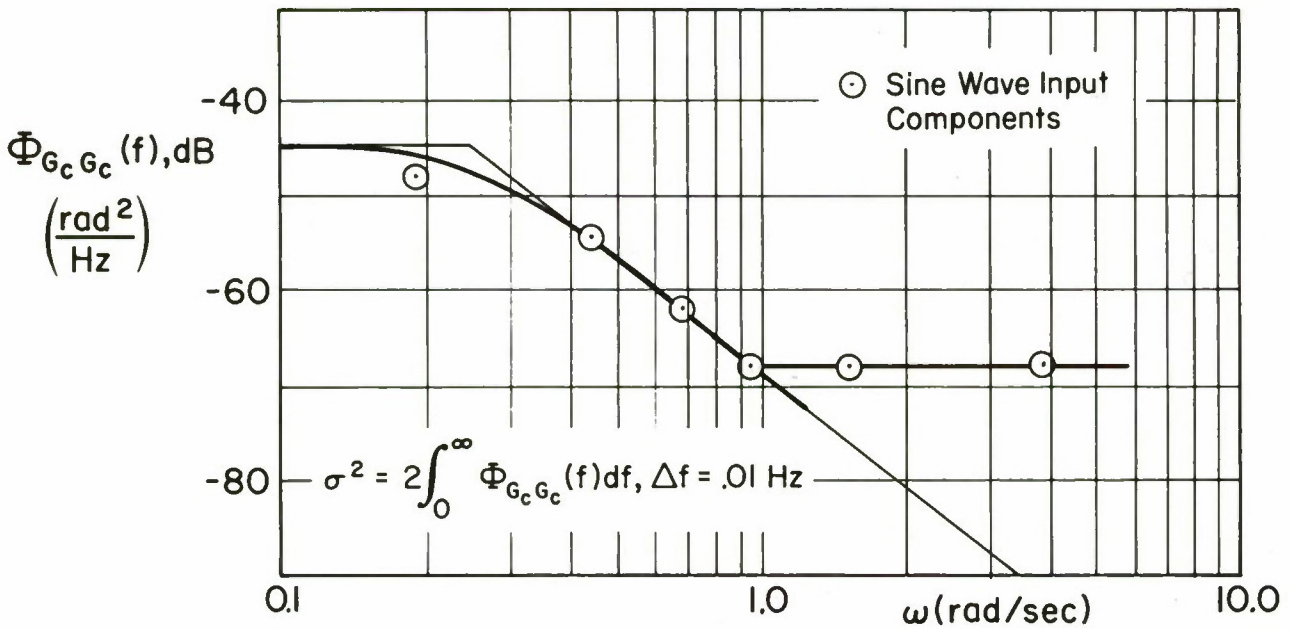


Figure 5. Power Spectral Density of Glide Slope Forcing Function

Prior to any simulator flying, each of the pilot subjects was given an overall briefing on the task and the equipment to be used. He was told to "try to keep the glide slope and localizer needles centered at all times." A data recording session usually involved 5 or 6 100 sec runs in succession, divided at random between manual ILS and flight director configurations. Fixed range and varying range configurations were not mixed in the same session, but were run on separate days.

PILOT SUBJECTS

Seven participated in the program, and data for four of them have been analyzed. The subjects were all current professional airline pilots or copilots, and reflected a cross section of age and background. Pilot 1 was a senior instructor captain with multiengine piston and jet bomber experience. Pilot 2, a younger copilot, transitioned into commercial flying from general aviation/light aircraft. Pilot 3, a young copilot, had extensive prior military single engine fighter/bomber experience; and Pilot 4 was a multiengine test pilot of long standing. Pilots 1, 2, and 4 had experienced this particular simulator in other research programs, but under slightly different instrument arrangements and test conditions. Pilot 4 was used in the shakedown runs associated with setting up and validating the simulation. Pilot 3 was unfamiliar with this simulator but had participated in previous STI experiments at other facilities. The simulation was sufficiently similar to their current experience that the pilots were able to achieve a stable level of proficiency within a few runs.

SCANNING TRAFFIC DEFINITIONS

Some definitions of the properties of the raw and reduced EPR data are needed. For a given run of T_R sec duration:

- M is the number of instruments
- N_i is the number of fixations on instrument i
- N_M is the total number of fixations on all instruments
- N is the total number of fixations on instruments, elsewhere, blinks, etc.

It follows that

$$N_M = \sum_{i=1}^{M} N_i$$

The duration of a look at a given instrument is called the "dwell time," T_d, and

Td_{i_k} is the duration of the kth dwell on instrument i

$$T_i = \sum_{k=1}^{N_i} Td_{i_k} \quad \text{is the total time fixating i}$$

$$T_R = \sum_{i=1}^{M} T_i + T_{other}$$

where T_{other} includes blinks and looks elsewhere than at the defined instruments. For data reduction convenience we assign a number to blinks and other regions of the panel so that all time during the run is subscripted and allocated.

Average properties of the data are important. The mean dwell time on instrument i is

$$\overline{T}_{d_i} = \frac{1}{N_i} \sum_{k=1}^{N_i} Td_{i_k} = \frac{T_i}{N_i}$$

The "scan rate" over all instruments on the panel is the average number of fixations per second, given by

$$\bar{f}_s = \frac{N}{T_R}$$

The scan rate on a given instrument is called the "look rate," given by

$$\bar{f}_{s_i} = \frac{N_i}{T_R}$$

The fraction of fixations on the ith instrument, ν_i, is called the "look fraction,"

$$\nu_i = \frac{N_i}{N}$$

The "dwell fraction" is the fraction of time spent on instrument i, given by

$$\eta_i = \frac{T_i}{T_R}$$

This is also called the "fractional scanning workload." The "look interval" is the inverse of the scanning workload, i.e.,

$$\bar{T}_{s_i} = \frac{1}{\bar{f}_{s_i}}$$

The look interval is a measure of the recycle time, and it can also be computed from the individual scan intervals (the time between successive looks at an instrument).

DATA REDUCTION PROCEDURE

The six instruments and other regions of the panel were numbered for analysis as shown in Fig. 6. Looks at region 8 were usually "blinks," and they resulted in the total workload on the instruments being less than unity. There were essentially no looks at regions 7, 9, and 10 in the data analyzed.

The eye point of regard is defined by vertical and horizontal coordinates on the panel. A typical segment of data is shown in Fig. 7. The major part

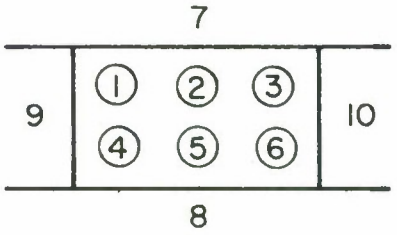


Figure 6. EPR Regions on Instrument Panel

of the dwell is well defined. The transitions between dwells take a small amount of time and may occasionally contain artifacts such as an overshoot, cross talk and fake looks (e.g., going from 5 to 1, passing over 2 but not dwelling on it). The transitions are defined as having a duration no greater than .15 sec. Typical vertical transition times (over all pilots) between instruments are in the range .06 to .09 sec. The horizontal transitions are slightly faster, .05 to .08 sec. The difference probably reflects eyelid lag on the vertical channel, which varies between subjects. If the transition times are longer than .15 sec they become an actual look, blink, etc. The transition times are allocated to the adjacent dwells in roughly equal proportions as shown in Fig. 7. An alternative procedure, deleting the transition times from the run, involves reduction difficulties that are avoided by the method used. Detailed artifacts of this type of EPR data such as scanning

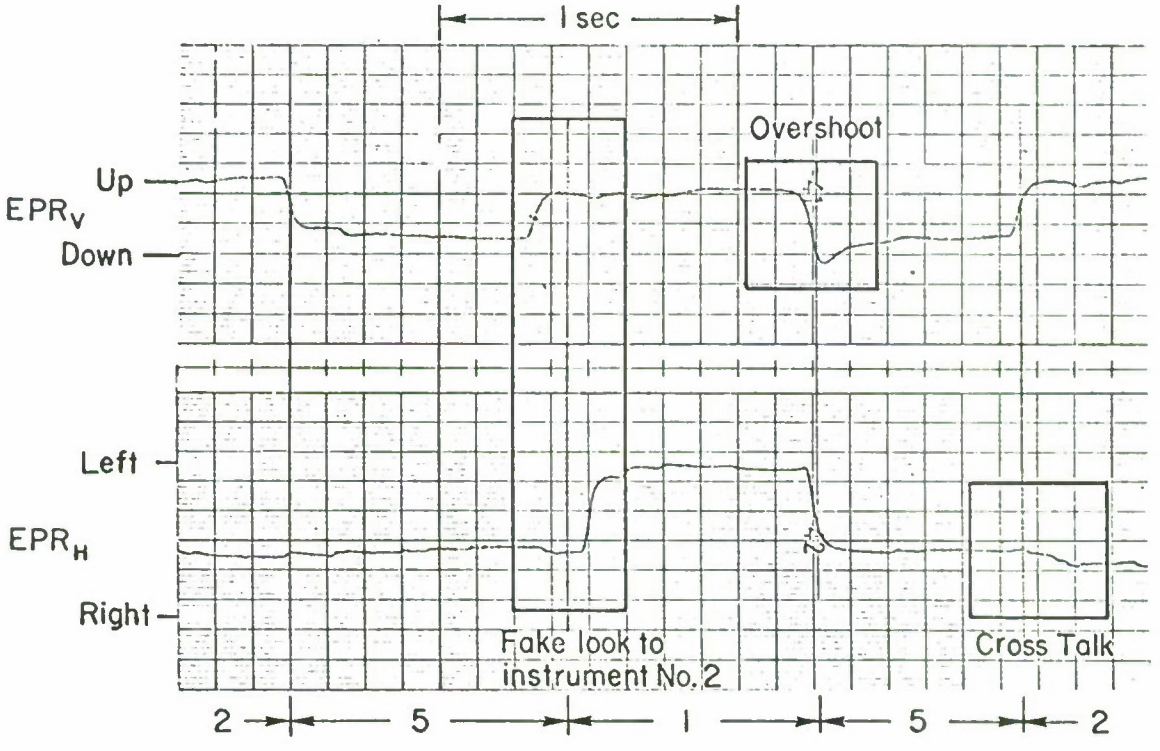


Figure 7. Illustrative EPR Data Sample

within a given instrument and blinks are discussed in Ref. 3. A Fortran IV program was written to statistically reduce the EPR data. The output consisted of the dwell time statistics defined above, histograms for each instrument, summations for all instruments, and one way link transitions between instruments.

SCANNING STATISTICS

Over 100 EPR data runs were made, involving a total of seven pilot subjects and six configurations (A through F). A cross section of 31 runs representing 3 subjects were selected for detailed analysis. The selected runs are listed in Table II. Each cell is denoted by a shorthand notation, e.g., C1 is Configuration C with Pilot 1. Configuration A is not included, because it was a single axis task with no scanning. Pilot 1 was the principal subject and replications for each configuration are shown. Runs for Pilots 2 and 3 help define interpilot and interconfiguration differences. Some of the range-varying runs involved "visual breakout" and transition to an outside visual field display. This transition occurred at least 10 to 15 sec after the 100 sec run interval, on verbal instructions from a "copilot." Analyses showed that it did not affect the pre-breakout data.

TABLE II

SUMMARY OF RUNS ANALYZED

CONFIGURATION	SUBJECT		
	1	2	3
B (Split-axis manual ILS fixed range)	B1 Two runs	B2 One run	
C (Manual ILS, fixed range)	C1 Five runs	C2 Five runs	
D (Manual ILS, varying range)	D1 Five runs		D3 Two runs
E (FD, fixed range)	E1 Four runs	E2 Three runs	
F (FD, varying range)	F1 Three runs		F3 One run

The detailed scanning statistics for each of the 31 runs were averaged over a given pilot/configuration cell to obtain Table III. The look rates, $\bar{f}_{s_i}$, and dwell times, $\bar{T}_{d_i}$, in Table III were examined to determine similarities and differences among the pilots and configurations. The results of these statistical tests and other observations are discussed below.

TABLE III

AVERAGE SCANNING STATISTICS

CONF.	SUBJ.	NO. RUNS AVG.	INSTRUMENT 1, IAS						INSTRUMENT 2, ATT OR ATT/FD					
			N_1	$\bar{T}_{d_1}$	σ_{T_1}	$\bar{f}_{s_1}$	η_1	ν_1	N_2	$\bar{T}_{d_2}$	σ_{T_2}	$\bar{f}_{s_2}$	η_2	ν_2
B	1	2	18	.57	.18	.09	.051	.068	89	.67	.29	.446	.296	.337
B	2	1	2	.64	.05	.02	.013	.015	61	.82	.36	.607	.498	.462
C	1	5	25	.69	.28	.05	.034	.042	238	.74	.33	.474	.353	.397
C	2	5	1	.63	0	.002	.001	.002	254	.95	.33	.511	.483	.453
D	1	5	43	.69	.35	.085	.059	.066	282	.81	.38	.560	.452	.431
D	3	2	3	.88	.40	.015	.013	.013	85	1.00	.45	.423	.425	.357
E	1	4	44	.55	.15	.110	.06	.113	171	1.73	1.72	.428	.739	.438
E	2	3	13	.54	.11	.043	.023	.070	84	3.03	3.41	.276	.836	.454
F	1	3	53	.56	.15	.177	.099	.163	147	1.50	1.36	.492	.737	.453
F	3	1	6	.59	.10	.059	.035	.079	37	2.19	1.80	.366	.80	.487

CONF.	SUBJ.	NO. RUNS AVG.	INSTRUMENT 3, PALT						INSTRUMENT 5, ESI/CSD					
			N_3	$\bar{T}_{d_3}$	σ_{T_3}	$\bar{f}_{s_3}$	η_3	ν_3	N_5	$\bar{T}_{d_5}$	σ_{T_5}	$\bar{f}_{s_5}$	η_5	ν_5
B	1	2	23	.38	.12	.115	.044	.087	111	1.02	.80	.557	.567	.421
B	2	1	2	.57	.04	.02	.011	.015	57	.78	.25	.567	.443	.432
C	1	5	40	.42	.16	.08	.034	.067	255	1.09	.72	.508	.551	.425
C	2	5	3	.41	.15	.01	.004	.009	242	.98	.40	.486	.476	.431
D	1	5	27	.45	.22	.054	.024	.041	265	.84	.43	.527	.444	.405
D	3	2	9	.44	.15	.045	.020	.038	96	.93	.51	.479	.443	.403
E	1	4	31	.40	.10	.128	.051	.131	76	.57	.20	.190	.109	.195
E	2	3	18	.50	.11	.059	.03	.097	41	.51	.12	.135	.069	.222
F	1	3	34	.35	.07	.113	.039	.105	59	.46	.13	.197	.090	.181
F	3	1	4	.34	.05	.04	.013	.053	27	.54	.25	.267	.145	.355

CONF.	SUBJ.	NO. RUNS AVG.	INSTRUMENT 6, IVSI						ALL INSTRUMENTS				
			N_6	$\bar{T}_{d_6}$	σ_{T_6}	$\bar{f}_{s_6}$	η_6	ν_6		N_M	$\sum T_1$	$\bar{f}_s$	
B	1	2	10	.40	.14	.05	.020	.038		254	200	1.32	
B	2	1	1	.55	0	.01	.005	.008		132	101	1.31	
C	1	5	16	.38	.07	.032	.012	.027		600	502	1.19	
C	2	5	2	.47	.23	.004	.002	.004		561	497	1.13	
D	1	5	5	.40	.18	.010	.004	.008		654	503	1.30	
D	3	2	43	.45	.14	.214	.095	.181		238	200	1.19	
E	1	4	16	.44	.11	.04	.018	.041		390	400	.976	
E	2	3	14	.49	.10	.046	.023	.076		185	305	.607	
F	1	3	9	.35	.12	.033	.012	.031		325	300	1.08	
F	3	1	1	.49	0	.01	.005	.013		76	101	.752	

Stationarity Within a Run

A key question in computing average scanning statistics during landing approach is whether there is any significant change in the pilot's scanning behavior as the approach progresses. One potential source of nonstationarity arises in the glide slope deviation bar and pitch flight director whose gains change in the range-varying configurations. Several of the range-varying runs were processed in three successive intervals and their statistics compared, in order to determine if the scanning statistics were nonstationary.

Table IV shows the mean dwell time and dwell fraction for each of the five instruments for three typical 100 sec runs. In general, the mean dwell times

TABLE IV

SCANNING STATISTICS FOR RUN SEGMENTS

INSTRUMENT	TIME INTERVAL (SEC)	RUN 26-05, D1		RUN 26-09, D1		RUN 26-07, F1	
		$\bar{T}_d$	η	$\bar{T}_d$	η	$\bar{T}_d$	η
1, IAS							
	0-33	.67	.080	.61	.056	.47	.116
	33-67	.56	.050	.66	.079	.65	.110
	67-100	.54	.082	.50	.044	.54	.102
2, ATT/FD							
	0-33	.78	.468	1.02	.557	1.38	.716
	33-67	.84	.529	.84	.477	1.64	.735
	67-100	.71	.500	.66	.440	1.25	.709
3, PALT							
	0-33	.53	.031	.47	.014	.40	.048
	33-67	.24	.007	.44	.026	.36	.040
	67-100	.50	.046	.39	.034	.33	.063
5, HSI/GSD							
	0-33	.75	.379	.68	.332	.46	.099
	33-67	.69	.395	.83	.395	.44	.086
	67-100	.68	.354	.85	.486	.43	.109
6, IVSI							
	0-33	.32	.010	.32	.010	.20	.006
	33-67	.61	.018	.26	.023	.40	.011
	67-100	0	0	.55	.016	.37	.012

do not change significantly* for successive thirds of runs. The dwell time on instrument 2 in the last one-third of Run 26-09 is significantly smaller than during the first two-thirds because the first 2 parts each have one very long dwell (i.e., 2.5 sec). If these long dwells are deleted there is no longer a significant difference. A comparison of Run 26-05 and 26-09 dwell times shows no significant differences for instruments 2 and 5. Other instrument dwell times were not compared in detail, because the number of looks were few. The tests for significance were not performed on the dwell fractions, but these data do not show any consistent trends between runs and their variability is probably not significant. Thus, these stationarity tests show no important range-varying effects and the EPR can be considered statistically stationary within a run.

Look and Scan Rates

Look rates involve the scanning frequency on a given instrument while scan rates involve the entire panel. Analyses of the results in Table III show that:

- The attitude or attitude/flight director instrument (No. 2) look rates are generally the same over all pilots and configurations.
- The HSI/GSD instrument (No. 5) look rates are significantly lower for flight director than manual ILS runs.
- The peripheral instrument look rates are scattered and show no strong trends.
- The all-instrument scan rates are significantly greater for the manual ILS runs than for the flight director runs.

These results correlate with the dwell time and workload results discussed below.

Peripheral Instrument Dwell Times

The dwell time is the average length of one instrument fixation. The peripheral instrument results in Table III show that:

- Mean dwell times on the altimeter instruments (No. 3) and IVSI (No. 6) are homogeneous over all pilots and configurations, and are not significantly different from one another. The mean is .42 sec.

---

*All tests of significance were performed at the 95 percent confidence level.

- The mean dwell times on the airspeed instrument (No. 1) are homogeneous over all pilots and configurations, and their average (.64 sec) is significantly greater than for the other peripheral instruments (3 and 6).

Some of the manual ILS data were quantized into .05 sec intervals to determine if there was any tendency for discretization in the dwell times (bunched at discrete durations). This did not occur.

Primary Instrument Dwell Times

The dwell time results on the attitude gyro (No. 2) and the HSI/GSD (No. 5) for the various pilots and configurations show that:

- Mean dwell times on the attitude and HSI/GSD instruments for Configuration B are often different from C and D, indicating that the additional lateral axes of control have an effect with some pilots. Note that bank angle is on 2 and localizer deviation and heading are on 5.

- Differences between fixed and varying range had no effect on the attitude instrument dwell times for either the manual ILS or flight director configurations. Intrapilot data on similar configurations can be lumped.

- Fixed versus varying range had a significant effect on the HSI/GSD instrument dwell times for both manual ILS and flight director configurations. In each case the varying range version had a shorter mean dwell time.

- The mean dwell times on attitude instrument with the manual ILS configurations are less than with the flight director at a very high level of significance. The dwell time variances are also much less.

- The dwell times on the HSI/GSD instrument show the opposite trend. The manual ILS means are greater than the flight director means at a high level of significance, as are the dwell time variances. These results are consistent, because attitude and localizer/glide slope are primary and share 80-90 percent of the scanning workload.

- Interpilot differences in mean dwell time on the primary instruments often occurred. These were most pronounced on the attitude instrument with the manual ILS configuration, and did not occur at all on HSI/GSD instrument with the flight director operating.

Fractional Scanning Workload

The dwell fraction, η_i, also called the fractional scanning workload, is the fraction of time during a run that the pilot is looking at that instrument. Average values for each instrument with each subject/configuration are given in Table III. Tests of significance were not made but certain trends are apparent:

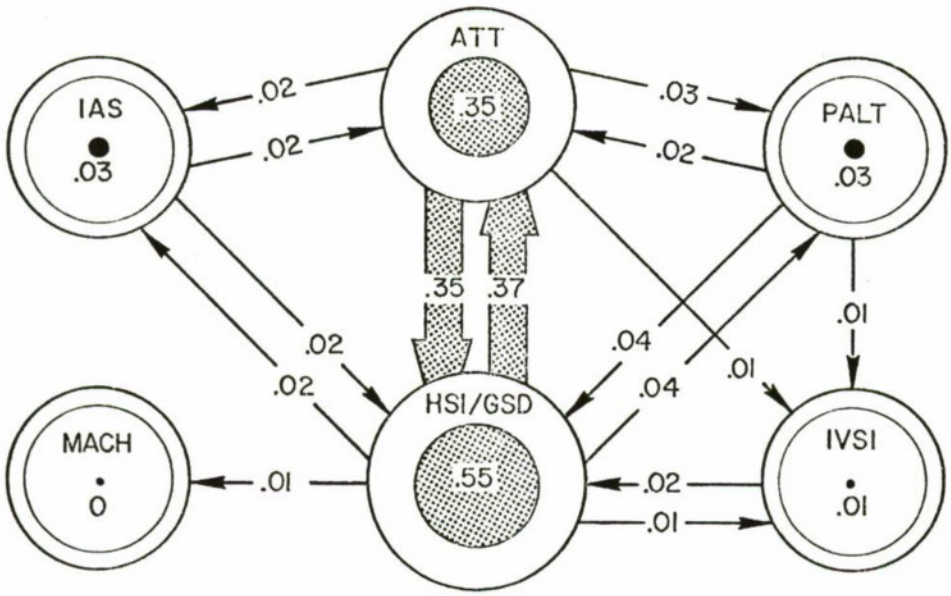
- The dwell fraction on peripheral instruments varies from run to run but there are no clear differences between pilots or configurations.
- The dwell fraction on peripheral instruments is much less than that on the primary instruments (by definition!).
- The dwell fractions on the attitude and HSI/GSD instruments are about equal with the manual ILS configurations.
- The dwell fraction on the attitude instrument is much larger with the flight director configurations than with the manual ILS ones.
- The dwell fraction on the HSI/GSD instrument goes way down when the flight director is in use, and it becomes effectively a peripheral instrument.

These differences in scanning workload are due mainly to differences in dwell time, and to some extent changes in scan rate as shown in Table III. For example, the unusually low E2 scan rate combines with the unusually long mean dwell time on the attitude instrument to give the highest observed mean scanning workload.

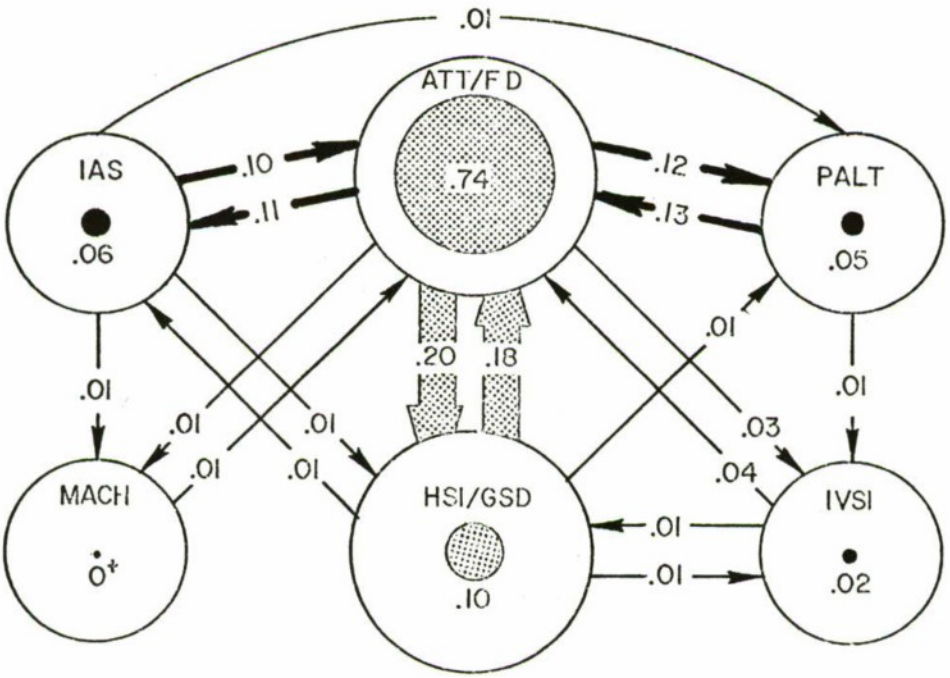
Transition Links

The one way link value, q_{ij}, is the fraction of all fixation transitions that go from instrument i to j; and conversely for q_{ji}. Their sum, $q_{ij} + q_{ji}$, is the two way link value and represents the fixation transitions between points i and j. Fitts, Jones, and Milton (Ref. 4) hypothesized that the link values between instruments are indicative of the goodness of panel arrangements. If the pilot is stationary over a run, one way link values are also indicative of dominant scan patterns.

The data have been reduced to show one way link values, and to determine the one and two way differences. Typical link vectors for Pilot 1 in the manual ILS and flight director tasks are compared in Fig. 8. The width of the link vector represents its magnitude, and the diameters of the shaded instrument centroids represent the dwell fraction. The sum of the dwell fractions is less than one due to blinks. The data show no dominant "circulation" of scanning, and the one way link values (e.g., 2 to 5 and 5 to 2)



Manual ILS Configuration, CI



Flight Director Configuration, EI

Figure 8. Typical Transition Link Vectors and Dwell Fractions

are approximately equal. The flight director task has a more evenly distributed percentage of scans to secondary instruments, although a high percent of time was spent on the attitude/flight director indicator. There were very few (i.e., < 1 percent) link transitions across instruments indicating the primary instruments were centrally located.

The big difference between the manual ILS and flight director configurations is related to the difference in scanning workload. The manual ILS results have dominant links from attitude to HSI/GSD and back. The flight director links appear more "scattered," because these central links are relatively smaller. Since the one way link vectors are about equal in opposite directions, the two way links can be used. This is closely related to the observation that the scan patterns show no strong evidence of circulation.

COMPARISONS WITH OTHER EYE SCANNING DATA

These eye scanning data are generally consistent with the results of prior research, where meaningful comparisons can be made. The largest and most thorough data on pilot eye movements were collected by Fitts, Milton, Jones, McIntosh, and Cole in a continuing program from 1949 to 1952, using in-flight eye camera films of 40 subjects. Of the published material, five (Refs. 4-8) were concerned with the landing approach phase. Similar, more recent, measurements have been made by Senders, Carbonell, and Ward in Ref. 12, utilizing electro-oculograms (EOG) of three subjects in a fixed base simulator. Their panel arrangement was identical to the standard instrument arrangement used by Fitts, et al., in Refs. 4 and 5 and shown in Fig. 9. An "experimental" panel used by Fitts, et al., in Refs. 6 and 7 is shown in Fig. 10 and more nearly resembles the panel arrangement used in this program.

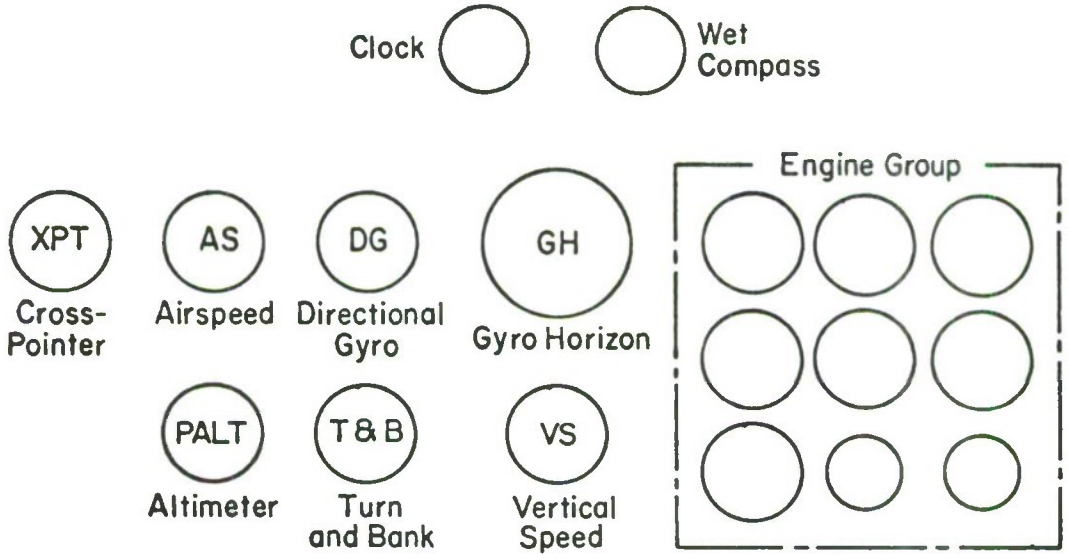


Figure 9. Standard Instrument Arrangement Used by Fitts, et al., (Refs. 4 and 5) and Senders, et al. (Ref. 12)

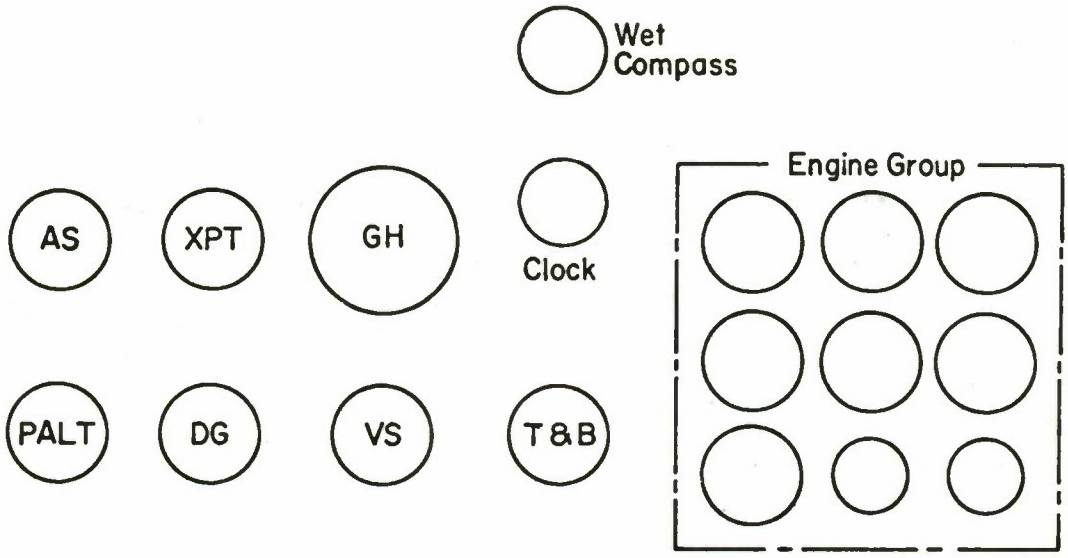


Figure 10. Experimental Instrument Arrangement
Used by Fitts, et al. (Refs. 6 and 7)

The last report of the Fitts series (Ref. 8) was concerned with flight director (actually zero reader) approaches. Measurements were made of 10 pilots who each flew one approach from the rear seat of a T-33 aircraft. The flight director indicator was a separate instrument as shown in Fig. 11. With this exception, the panel was similar to the experimental instrument arrangement of Refs. 6 and 7.

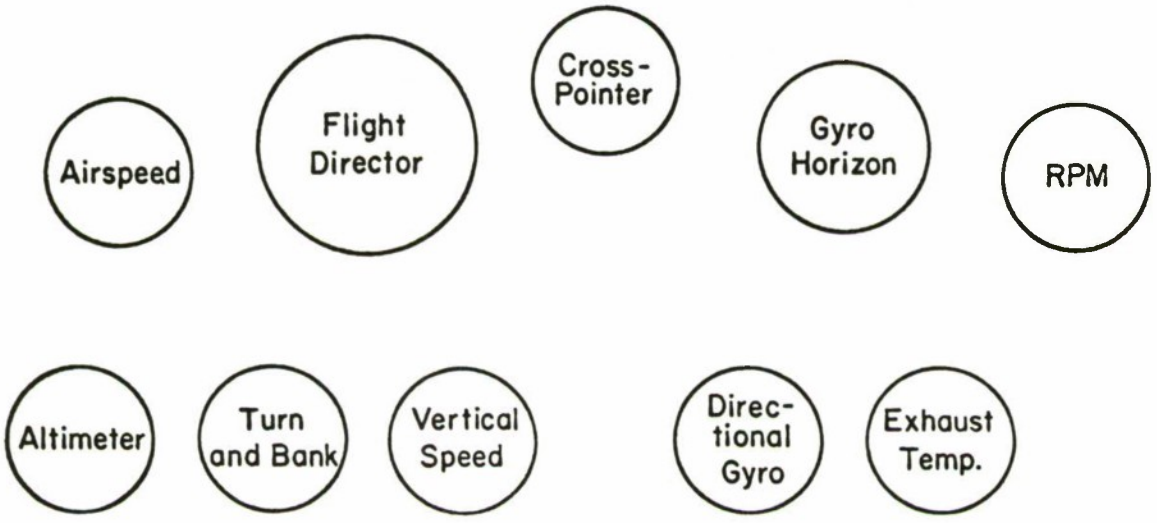


Figure 11. Panel Arrangement Used by Fitts, et al. (Ref. 8)
for Flight Director Approaches

Table V compares dwell fractions from Fitts, et al. (Refs. 4 and 6) and Senders, et al. (Ref. 12) with the manual ILS data from the DC-8 simulation. The individual dwell fractions of the directional gyro display and the cross-pointer display for the past data were summed to compare with the dwell fractions for the integrated HSI/GSD display. The sum of the dwell fractions

TABLE V

COMPARISON OF DWELL FRACTIONS WITH PAST DATA

INSTRUMENT	DWELL FRACTION			
	STI*	SENDERS† (REF. 12)	FITTS† (REF. 4)	FITTS† (REF. 6)
ATTITUDE	.438	.282	.15	.11
HSI/GSD	.472	.445	.655	.650
PALT	.018	.070	.02	.02
AS	.020	.073	.10	.07
IVSI	.035	.128	.05	.05

*Average of 3 pilots (Configurations C and D)

†Average of 2 pilots (Phase III)

†Average of 40 pilots (instrument low-approaches)

(workload) on instruments in these tables is less than unity; due to blinks (and looks at 4) and because other (noncomparable) instruments were present in the other data. The comparison of available mean dwell times is presented in Table VI.

The two way link values for the weighted average of the three pilots in the manual ILS configurations (C and D) are shown in Table VII along with the experimental panel results of Fitts, et al. (Ref. 6). Since the number of instruments differed a direct comparison is not meaningful; however, the primary link in the Fitts data is XPT-DG which is the currently integrated HSI/GSD display. Combining the second largest link, XPT-GH, with the XPT-DG would approach the current HSI/GSD-ATT link value.

Tables V, VI, and VII serve to compare rather than evaluate the data. There are many differing factors in the three sets of data which would influence the results. For example, the gyro horizon was just replacing the "needle-ball-airspeed" technique during the period of Fitts' studies. The current technique supported by nearly all pilots rests primarily on attitude control and therefore has a higher fractional workload and associated dwell time. The peripheral instruments (i.e., PALT, IVSI) are comparable and do show similar workloads and dwell times.

TABLE VI

COMPARISON OF MEAN DWELL TIMES WITH PAST DATA

INSTRUMENT	MEAN DWELL TIME (SEC)	
	STI*	FITTS (REF. 6)
ATTITUDE	.85	.37
HSI/GSD	.96	XPT .76 DG .54
PALT	.43	.38
AS	.70	.49
IVSI	.43	.39

*Weighted average of 3 pilots (Configurations C and D)

TABLE VII

COMPARISON OF AVERAGED TWO WAY LINK VALUES

STI		FITTS (REF. 6)	
LINK	VALUE	LINK	VALUE
HSI/GSD-AS	.03	XPT-AS	.10
HSI/GSD-ATT	.77	XPT-GH	.22
HSI/GSD-PALT	.04	XPT-PALT	.02
HSI/GSD-IVSI	.04	XPT-VS	.04
ATT-AS	.04	XPT-DG	.31
ATT-PALT	.04	GH-AS	.02
ATT-IVSI	.02	GH-ALT	<.02
AS-IVSI	0	GH-VS	.02
AS-ALT	0	AS-PALT	.02
		AS-DG	.02
		GH-DG	.06
		VS-DG	.05
		PALT-DG	.02

The theory (Ref. 1) indicates that scanning rates should be strongly influenced by the vehicle dynamic properties and so should the workload; because (for the same dwell time) the scanning workload will be directly proportional to the scan rate (i.e., $\omega_{si} = \overline{T_{di} f_{si}}$). Additional factors noted by Fitts, et al., as having a significant effect on the dwell time and fraction were interpilot differences, day versus night operation, and manual ILS versus ground controlled approaches (GCA).

The flight director data of Ref. 8 are shown in Table VIII with those obtained from the DC-8 simulation (Configurations E and F combined). The dwell fractions (fractional scanning workload) can be added and as such compare very closely with the current attitude/flight director result. Mean dwell times, a parameter less dependent on vehicle properties, exhibits the

TABLE VIII

COMPARISON OF FLIGHT DIRECTOR EYE TRAFFIC WITH PAST DATA

INSTRUMENT	DWELL FRACTION		MEAN DWELL TIME (SEC)		SCAN RATE (LOOKS/SEC)	
	STI	FITTS (REF. 8)	STI	FITTS (REF. 8)	STI	FITTS (REF. 8)
[FLIGHT DIRECTOR] ATTITUDE	.77	.64 .13	1.94	1.29 .48	.40	.50 .27
[LOCALIZER/ GLIDE SLOPE] HEADING	.095	0 .01	.52	.25 .50	.18	.022 .02
AIRSPEED	.055	.09	.55	.52	.11	.168
VERTICAL SPEED	.02	.02	.44	.45	.04	.049
ALTIMETER	.04	.01	.40	.42	.10	.037
MISC.	.02	.10	—	—	.06	.46

same trends between instruments and may have an additive property for integrated displays. The total scan rate in the Ref. 8 study was 92/min or 1.53/sec. This compares well to the .89/sec in the DC-8 study. The scan rates show that the flight director is dominant. The large number of miscellaneous scans in the Ref. 8 data was due mainly to unresolved looks, and also included looks at the rpm, exhaust temperature, and turn and bank indicators, blinks and looks at switches.

The link values recorded in Ref. 8 support the finding that the one way links between pairs of instruments are approximately equal, and that the flight director is the center of attention. Table IX presents the link values between pairs of instruments, disregarding the values less than 2 percent. The STI data are for Configurations E and F combined.

TABLE IX

COMPARISON OF AVERAGED TWO-WAY LINK VALUES WITH PAST DATA WITH FLIGHT DIRECTOR DISPLAYS

INSTRUMENT LINKS	STI	FITTS (REF. 8)
FD-ATT	Integrated	.38
FD-AS	.24	.24
FD-PALT	.21	.04
FD-VS	.07	.05
FD-RPM	0	.06
ATT-RPM	0	.03
FD-T&B	0	.03
FD-HSI/GSD	.41	<.02

Thus, although most of the Fitts, et al., and Senders, et al., data were taken with a panel which differs somewhat from the current T-layout, their results can be compared with the new non-flight director data. The STI data show a larger dwell fraction on the gyro horizon than past data. The dwell fraction for the HSI/GSD instrument in the STI data is less than the lumped dwell fraction for directional gyro and crosspointer in either the Fitts or Senders results. There are only minor differences in other dwell fractions. The dwell times are similar among respective peripheral instruments and between the then-primary and the current primary displays. Differences in dwell fraction can be attributed to differences in panel arrangement, and to the evolution of a pilot technique using attitude control. The Fitts zero reader dwell fraction data (Ref. 8) agree well with the present (flight director configuration) data for most instruments. The exception is that the STI data show a larger dwell fraction on HSI/GSD than does the sum of his crosspointer and DG data. The respective dwell times are comparable.

In summary, apparatus is now available which allows the measurement and rapid reduction of eye-point-of-regard data for realistic vehicle control tasks. The results of recent simulation studies provide new insight into pilot scanning behavior for instrument approach; and, when combined with concurrently measured pilot response data, will serve to validate and extend the theory of manual control displays. The current data are consistent with past data for comparable flight control tasks.

REFERENCES

1. McRuer, Duane, Henry R. Jex, Warren F. Clement, and Dunstan Graham, Development of a Systems Analysis Theory of Manual Control Displays, Systems Technology, Inc., Tech. Rept. 163-1, Oct. 1967.

2. 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., Tech. Rept. 170-2, July 1969.

3. Weir, David H., and Richard H. Klein, The Measurement and Analysis of Pilot Scanning and Control Behavior During Simulated Instrument Approaches, Systems Technology, Inc., Tech. Rept. 170-4, June 1969.

4. Milton, J. L., R. E. Jones, and P. M. Fitts, Eye Fixations of Aircraft Pilots: Frequency, Duration, and Sequence of Fixations When Flying the USAF Instrument Low Approach Systems (ILAS), USAF Tech. Rept. 5839, Oct. 1949.

5. Fitts, P. M., R. E. Jones, and J. L. Milton, Eye Fixations of Aircraft Pilots: Frequency, Duration, and Sequence of Fixations When Flying Air Force Ground Controlled Approach System (GCA), AF Tech. Rept. 5967, Feb. 1950.

6. Milton, J. L., B. B. McIntosh, and E. L. Cole, Eye Fixations of Aircraft Pilots: Fixations During Day and Night ILAS Approaches Using an Experimental Instrument Panel Arrangement, USAF Tech. Rept. 6570, Oct. 1951.

7. Milton, J. L., B. B. McIntosh, and E. L. Cole, Eye Fixations on Aircraft Pilots: Fixations During Day and Night GCA Approaches Using an Experimental Instrument Panel Arrangement, USAF Tech. Rept. 6709, Feb. 1952.

8. Milton, J. L., and F. J. Wolfe, Eye Fixations of Aircraft Pilots: Fixations During Zero Reader Approaches in a Jet Aircraft, WADC Tech. Rept. 52-17, Feb. 1952.

9. Gainer, C. A., R. D. Monroe, J. E. Brown, et al., All Weather Landing Simulation for Category III Airborne Configuration: Volume I. Summary of Studies on Flight Directors and Split Axis Control, Bunker-Ramo Corp., July 1967.

10. Watts, A. F. A., and H. C. Wiltshire, Investigation of Eye Movements of an Aircraft Pilot under Blind Approach Conditions, College Note No. 26, The College of Aeronautics, Cranfield, May 1955.

11. Lennox, D., Airline Pilots' Eye Movements During Take-off and Landing in Visual Meterological Conditions, Aeronautical Research Lab., Human Engineering Note 15, Australian Defense Scientific Service, Aug. 1963.

12. Senders, John W., Jaime R. Carbonell, and Jane L. Ward, Human Visual Sampling Processes: A Simulation Validation Study, NASA CR-1258, Jan. 1969.

13. Stapleford, R. L., D. T. McRuer, and R. E. Magdaleno, "Pilot Describing Function Measurements in a Multiloop Task," Second Annual NASA-University Conference on Manual Control, NASA SP-128, 1966, pp. 181-204.

14. Stapleford, Robert L., Samuel J. Craig, and Jean A. Tennant, Measurement of Pilot Describing Functions in Single-Controller Multiloop Tasks, NASA CR-1238, Jan. 1969.

TACTICAL UTILITY HELICOPTER INFORMATION TRANSFER STUDY

JOHN A. BARNES
Department of the Army
Human Engineering Laboratories
Aberdeen Proving Ground, Maryland

ABSTRACT

This study was an effort to determine what information, available from basic flight instruments, was used by the pilot to perform the various missions that encompass the Tactical Utility Helicopter Mission. A UH-1 aircraft, the current U. S. Army tactical utility helicopter, was used as the test vehicle and all measures relate to this vehicle.

Figure 1. EMC-2 Eye movement Camera System in flight use.

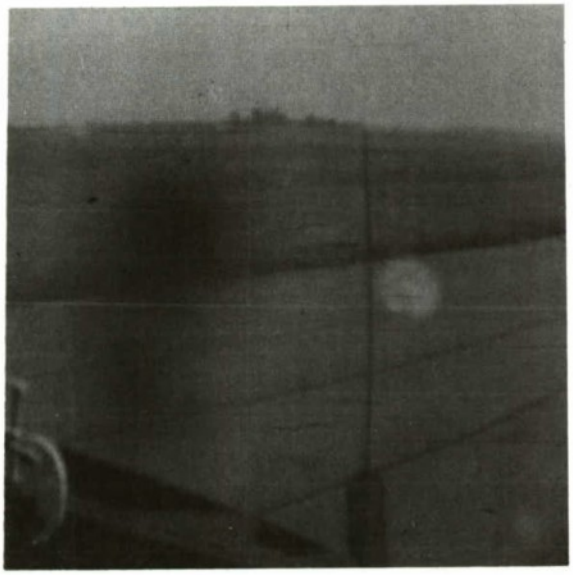


Figure 2. A fixation point during hover in ground effect.

Figure 3. A fixation point during instrument climb.

Figure 4. A fixation point during low level flight.

TACTICAL UTILITY HELICOPTER INFORMATION TRANSFER STUDY

This presentation is based on Department of the Army Human Engineering Laboratories Technical Memorandum 7-70 and provides a close look at the method used to determine what instruments were used by the pilots to perform the various segments of the test missions. The program considered the individual segments of three types of tactical utility helicopter missions; (a) Utility Transport, (b) Fire Support, (c) Rescue. These mission segments or tasks used the task descriptions given in U. S. Army Aviation Material Laboratories Technical Report 68-39, A Study of Handling Qualities Requirements of Winged Helicopters, as parameters. An example of the tasks are given in Figure 5. The test missions

TASK	DESCRIPTION
CLIMB FROM HOVER	Acceleration at least 3K per second, heading and roll attitude $\pm 5^{o}$, and rate of climb at least 500 feet per minute.
CLIMB, INSTRUMENTS	Airspeed $\pm$ 10K, heading $\pm 10^{o}$, rate of climb 500 feet per minute $\pm$ 100 feet per minute, and attitude $\pm 5^{o}$.
INITIAL DESCENT, VISUAL	At approach speed $\pm$ 10K maintain 30^{o} descent, $\pm 5^{o}$, attitude and heading $\pm 3^{o}$.
INITIAL DESCENT, INSTRUMENTS	At approach speed $\pm$ 5K maintain 15^{o} descent, $\pm 3^{o}$, attitude and heading $\pm 3^{o}$.

Figure 5. Tactical Utility Helicopter Tasks

were designed to allow the proper amount of time for the pilot to perform the complete task or to fly for at least two minutes performing a given task, whichever was shorter. The length of these missions, 20 minutes, was determined by the film capacity of the Westgate EMC - 2 eye movement camera system. This system was used to record the eye movements of the pilots while flying the aircraft and to determine the exact eye fixation point at all times during the flights and also to determine the scanning patterns

MANEUVER	START	END
Take Off	00:00	00:00
Hover IGE	00:00	00:02
Vertical Climb to 1000'	00:02	00:04
Cruise IFR	00:04	00:07
Standard Rate Turn IFR	00:07	00:08
Climb IFR to 1500'	00:08	00:09
Cruise IFR	00:09	00:12
180° Turn IFR	00:12	00:13
Steep Approach IFR to 500'	00:13	00:15
Hover OGE VFR	00:15	00:16
Vertical Descent	00:16	00:18
Land		00:19

Figure 6. Mission Plan I, flown at 60 knots or less.

MANEUVER	START	END
Take Off	00:00	
Climb IFR to 1500'	00:00	00:03
Cruise IFR	00:03	00:06
Standard Rate Turn	00:06	00:07
Cruise IFR	00:07	00:10
Descent IFR to 500'	00:10	00:12
Descending Turn IFR to 200'	00:12	00:13
360° Hovering Turn VFR	00:13	00:16
Vertical Descent	00:16	00:18
Land		00:19

Figure 7. Mission Plan II, flown at 100 knots or greater.

of the pilots. While it was possible to record the exact point upon which the pilot was fixating this study was only concerned with the instrument that was being looked at, not the point. The EMC-2 camera system had been modified by the addition of a fitted

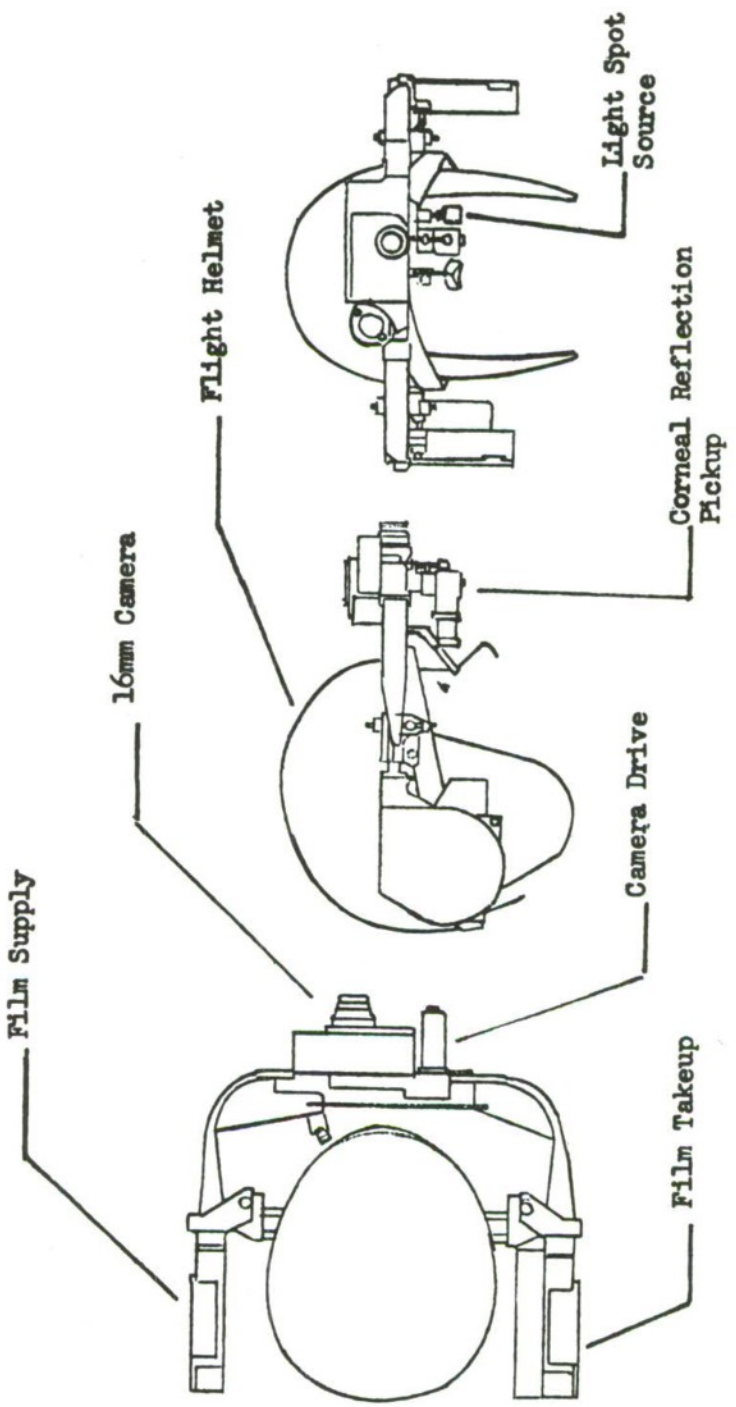


Figure 8. EMC-2 Camera System

TASK	INFORMATION DESIRED	PITCH ANGLE	ROLL ANGLE	ALTITUDE	AIRSPEED	STEERING	ANGLE OF ATTACK	VERTICAL VELOCITY	TURN RATE	HOVER POSITION	HOVER GROUND SPEED	HEADING	GROUND SPEED	TRACK	TORQUE	RPM	ENGINE CONDITION	EXTERNAL SOURCE	INTERNAL SOURCE
SPOT HOVER IGE		9	9	9						27	27	18	27		73	82	36	91	9
SPOT HOVER OGE		9	9	73	27	91		9			27	55	27	36	36	36	27	82	18
360° HOVERING TURN OGE		36	36	55	18	91			18		9	73	9	18	64	64	27	91	9
VERTICAL CLIMB		36	36	64		73		73			18	45	18	27	73	73	36	82	18
VERTICAL DESCENT		36	36	64	9	73		91				45		27	55	73	27	73	27
CRUISE, 60K, VISUAL		73	73	91	100	100		82	36			82		27	55	45	45	55	45
CRUISE, 60K, INSTRUMENTS		91	91	91	100	100		91	36			91		36	55	45	36		100
STANDARD RATE TURN, 60K		100	100	100	100	100		73	82			82		36	27	27	18		100
CLIMB, 60K, 500 FPM		73	73	91	82	91		100	27			73		18	55	27	9		100
CLIMB FROM HOVER		64	64	64	55	73		91	18			64		9	82	55	18		100
INITIAL DESCENT TO 500'		82	82	91	91	100	20	91	18		18	82	18	45	73	27	18	18	82
REVERSE DIR. OF FLIGHT, 60K		82	82	64	73	45		18	91			36		9	64	45	27	36	64
CRUISE, 100K, VISUAL		55	55	100	91	100		55	27			64		45	45	55	45	64	36
CRUISE, 100K, INSTRUMENTS		91	91	100	100	100		100				82		36	36	45	27		100
STANDARD RATE TURN, 100K		100	100	100	100	100		64	82			73		36	36	36	18		100
TERRAIN FOLLOWING, 100K		9	9	45	45	55		9	45			45		9	45	27	18	100	
CLIMB, 100K, 500 FPM		73	73	82	91	91		100	27			73		18	55	27	18	18	82
DESCENT, 100K, 500 FPM		73	73	82	91	100	33	91	27			91		36	73	36	18		100
180° DESCENDING TURN, 100K		91	91	100	73	100		91	100			82		27	45	27	18		100
REVERSE DIR. OF FLIGHT, 100K		82	82	82	82	55		27	100			36		18	73	64	36	45	55

TABLE 1. PERCENT OF PILOTS DESIRING SPECIFIC INFORMATION FOR TASK PERFORMANCE

TASK	INSTRUMENT USED	ATTITUDE INDICATOR	ALTITUDE	AIRSPEED	COMPASS	VERTICAL VELOCITY	TURN RATE	TORQUE	RPM	ENGINE INSTRUMENTS	EXTERNAL REFERENCE
SPOT HOVER IGE VISUAL		3	2	1	2	2	1	2	2	1	85
SPOT HOVER IGE INSTRUMENTS		25	12	7	18	12	3	15	13	11	
SPOT HOVER OGE		27	22	17	17	20	2	11	12	7	
360° HOVERING TURN OGE		15	19	9	1	5		5	5	3	38
VERTICAL CLIMB		29	30	9	7	20	3	7	6	5	
VERTICAL DESCENT		4	5	4	3	1		2	4	7	75
CRUISE, 60K, VISUAL		10	8	7	13	7	1	5	4	4	48
CRUISE, 60K, INSTRUMENTS		23	21	12	14	10	2	8	8	11	
STANDARD RATE TURN, 60K		23	18	9	19	9	7	6	6	6	
CLIMB, 60K, 500 FPM		19	17	11	19	19	2	8	6	7	
CLIMB FROM HOVER		23	23	7	6	16	3	6	5	4	15
INITIAL DESCENT TO 500', 60K		31	27	24	17	24		6	7	10	
REVERSE DIR. OF FLIGHT, 60K		23	18	9	19	9	7	6	6	6	
CRUISE, 100K, VISUAL		10	8	7	13	7	1	5	4	4	48
CRUISE, 100K, INSTRUMENTS		23	21	10	15	10	2	4	5	5	
STANDARD RATE TURN, 100K		14	25	2	25	14	1	3	3	3	
TERRAIN FOLLOWING, 100K		2	2	2	2	2		2	2	2	85
CLIMB, 100K, 500 FPM		8	15	13	10	12	3	4	9	3	
DESCENT, 100K, 500 FPM		27	29	16	11	17	2	6	7	7	
180° DESCENDING TURN, 100K		35	18	15	8	8	3	3	4	5	3
REVERSE DIR. OF FLIGHT, 100K		14	25	2	25	14	1	3	3	3	

TABLE 2. PERCENT OF TIME INSTRUMENT USED DURING FLIGHT

bite bar which insured that once calibrated for a flight the system would stay in calibration until removed from the subject. The unique feature of the bite bar attachment was the acrylic mouthpiece which was fitted to the subject and covered his teeth, upper and lower, from canine to canine. The average wearing time for the system was 30 minutes and there were no complaints of fatigue.

The results of these flight missions are shown in Table 2. These data were gathered from four flights, 80 minutes of flight time, flown by two pilots in a UH-1B. These men were highly experienced and had instrument ratings and combat experience. The numbers indicate the percentage of the total time available for the particular task that the pilot was fixating on the listed instrument. In cases where the sum of these numbers exceeds 100 the pilot was fixating on a point between two instruments and both were recorded as the instrument of interest. This condition occurred most frequently during climb and descent tasks and the point of fixation was between the Vertical Velocity Indicator and the Altimeter. This is an indication that the pilot was using parafoveal vision to attend to both of the instruments without moving his fixation point.

Prior to the flight phase of this study eleven pilots who were current in the UH-1 were presented with all of the tasks described in TR 68-39 and asked to indicate which instruments/information they required to perform the task. The results of this survey are presented in Table 1. The numbers represent the percentage of the pilots who felt that they required specific flight information to perform the task.

REFERENCES

1. Barnes, J. A., Tactical Utility Helicopter Information Transfer Study, Dept. of the Army, Human Engineering Laboratories, Aberdeen Proving Ground, Maryland. TM 7-70, April 1970.

2. Sardanowsky, W., Harper, H. P., A Study of Handling Qualities Requirements of Winged Helicopters, U. S. Army Aviation Material Laboratories, Fort Eustis, Virginia. TR 68-39.

A PERSPECTIVE GLIDESLOPE INDICATING SYSTEM

Noel A.J. Van Houtte

Man-Vehicle Laboratory
Department of Aeronautics and Astronautics
Massachusetts Institute of Technology

ABSTRACT

A perspective glideslope indicating system has been evaluated as an alternative to conventional aircraft displays for landing or contact analog displays.

A V/STOL aircraft (of the tilt-engine type) has been simulated in six degrees of freedom and has been flown from cruise altitude to touchdown with severe wind disturbances, descending along glideslopes of 4.5°, 8.8° and 17.3°. Distinct control techniques have been observed among three classes of pilots of different experience.

The value of the perspective glideslope display has been shown 1) in the ease of performing coordinated maneuvers, allowing large but quite precise changes of the flight variables, 2) in the consistency of touchdowns, 3) in accuracy of tracking the glideslope with dead beat response, 4) in the learning curve, and 5) in the effectiveness of the representation of the integrated real world outside picture.

The display is generated on a large screen cathode ray tube, using an analog line drawing scheme. The lines are derived from inertial coordinates of characteristic points of runway and glideslope relative to the aircraft, and are updated by a digital machine, using the translational and rotational rates of change, resulting from the simulated flight of the aircraft.

INTRODUCTION

On December 8, 1964 a fully automatic landing was performed at Wright-Patterson Air Field with a Caravelle of United Air Lines in a demonstration to Air Force Officials[1]. This event was described in many newspapers at that time. Since then, work has been continued in two different directions : the fully automatic landing on the one hand, and the manual control problem using electronic displays on the other. The latter one still keeps the pilot actively in the loop, and is the subject of this paper.

The systems configuration is shown in Figure 1. The pilot workload consists of three distinct parts : interpretation of the read-out of the instruments, the decision making process, and finally the control action to be taken. The display as

---

This research is supported by NASA Grant NsG-577

indicated can be a conventional set of instruments, but for our application it will be the electronic display. The latter group can be divided into two categories : integrated displays and contact analog displays [2]. The display that will be presented here is a contact analog display of the skeleton type. The rationale in setting up this format has been to reduce the pilot workload by combining the interpretation part with the display, and to let the computer do the 'think work'. However this is theory, and theory and practice are quite often two different things.

EXPERIMENTAL SET-UP

The vehicle used in this investigation is a V/STOL aircraft of the tilt-engine type [3]. Figure 2 shows the fixed base simulator used. The conventional controls about each of the three axes are indicated on the right-hand side, while the controls of the tilt-engines are explained on the left-hand side. Horizontal motion of the stick allows variation of the throttle setting while twisting the stick allows rotation of the engines. The little white bar indicates the thrust direction with respect to the aircraft.

Figure 3 shows the perspective display, drawn by an analog line drawing mechanism [4]. The location of the lines on the screen is derived from the end points, whose inertial position is determined by means of a digital computer. The lines selected for the picture are the runway lines, the glideslope lines as if there were a roadway in the sky, and vertical poles down to the ground. Poles are set at the beginning of the descent and also at 20,000 ft from the runway threshold.

A side view of this configuration is given on the bottom of Figure 4. The top of the figure shows five screen images of selected views along the glidepath. The first one shows the image at the beginning of the experiment, with the airplane trimmed in level flight. The second one shows a position of the airplane where the closest vertical poles are almost off-screen. The aircraft is above the extended glideslope lines, as indicated by the lines of the roadway which are bent downward from the center to the screen edge. The nose of the aircraft is aimed at the bottom of the vertical poles, set at 70,000 ft. The third one shows the aircraft little before the second set of vertical poles. The aircraft is now below the glideslope and this is indicated by the lines in the picture being bent upward from the center to the screen edge, as if one was looking at the underside of the roadway. The nose of the aircraft is aimed at a point on the ground, which is further than the far end of the runway. The fourth image shows the aircraft right on the glide-slope, and this is indicated by the glideslope lines in the image being horizontal. The nose of the aircraft is aimed at the far end of the runway. The last image shows the aircraft just before touchdown. Touchdown happens precisely when the runway lines on the screen become one straight line.

The simulated aircraft has been flown by eight subjects, namely three sets of pilots with different flight experience. The task is indicated on top of Figure 5. It consists of level flight for the first 10,000 ft, then the descent along the prescribed glideslope (4.47°, 8.87° and 17.3°) for the next period and finally transition, hover and landing. The performance was measured using a score factor which takes into account the touchdown error in range, the vertical velocity at touchdown and also the absolute error of the deviation of the prescribed glideslope. The latter error was taken into account using a weighting factor which decreases with the distance from the runway threshold. This performance score factor was then plotted versus the number of trials for each display, as in Figure 6, and this figure was used to make sure that the steady state level had been reached. This graph was also used in conjunction with later data reduction. The first two points on the plot are an indication what the steady state level will be, and are the points gotten without applying the wind disturbance.

BRIEF DISCUSSION OF THE RESULTS

Some of the results with the perspective display are compared with the conventional instruments in the next figures. Figure 7 shows the touchdown mean using the conventional instruments, averaged over all three glideslopes. Pilots one through three are inexperienced subjects. Pilots four and five have limited experience, while the pilots six through eight are well-trained pilots. The desired touchdown point was 90,000 ft and no special penalty was given for landing short. Figure 8 shows the same results for the perspective display. The standard deviation of the error has decreased and the mean is closer to where one expects it. Figure 9 shows the impact velocity using the conventional instruments. The values are quite high because of the steep descents imposed. Figure 10 shows the same results but for the perspective display and again improvement is observed.

It is hard to compared the results of figures 7 through 10 with numbers in real life. The reason for it is that the task was quite difficult. Indeed the simulated aircraft did not have augmented stabilization, and wind disturbance could reach a maximum of 50 ft/sec. (The wind amplitude was proportional to altitude however.) Furthermore the control of a V/STOL aircraft can be done in various ways, and is not a process that can be learned easily, since there is more than one main variable involved. In the present application there are three : pitch attitude, power setting and thrust direction.

Observation of the pilot describing function for tracking the glideslope revealed that there are different levels of control : gain, first order lead and second order lead. Inexperienced subjects begin their training by acting as a pure gain. Improvement in their performance brings them to the next category. Pilots with limited experience start exhibiting a

first order lead in their describing function. After some time they reach the third level of control strategy. The highly skilled pilots exhibit a second order lead. They achieve this by applying a control technique in a pulsatile manner. Different techniques were observed depending upon the glideslope. For the steep glideslopes, for example, the experienced pilots behave as a time-varying system, a combination of first and second order lead strategies. With the perspective display however, all classes of pilots become one group.

This can probably be best illustrated with the learning curve. Figure 6 shows definitely a learning period for the experienced pilot, using the dial instruments. However this learning effect is hardly noticeable in Figure 11, which is the learning curve for an inexperienced subject using the perspective display. Note too the order of magnitude difference between the score factors.

CONCLUSION

The unique feature of the display is that it restores the visual cues of the outside world in a skeleton fashion, but it has in addition the quantitative information necessary for the landing task.

It gives the attitude information, which is displayed by the most distant points of the visual field, as well as position information, which is given by the closest points displayed at the edge of the screen. Both states are visually linked in the picture, while this link is not so evident in the conventional instruments.

BIBLIOGRAPHY

1. Anon. 'Avionics', Electronics Review, Vol. 37 #32, December 28, 1964.

2. Ketchel, James & Jenney, Larry 'Electronic and Optically generated Aircraft Displays. A Study of Standardization Requirements', Janair report No. 680505, May 1968.

3. Van Houtte, Noel 'Display Instrumentation for V/STOL Aircraft in Landing', Sc.D. Thesis, Department of Aeronautics and Astronautics, M.I.T., June 1970.

4. Young, L.R. & Li, Y.T. 'Studies of Human Dynamic Space Orientation using Techniques of Control Theory', 7th semi--annual status report on NASA Grant NsG-577, Man-Vehicle Laboratory, M.I.T., June 1967.

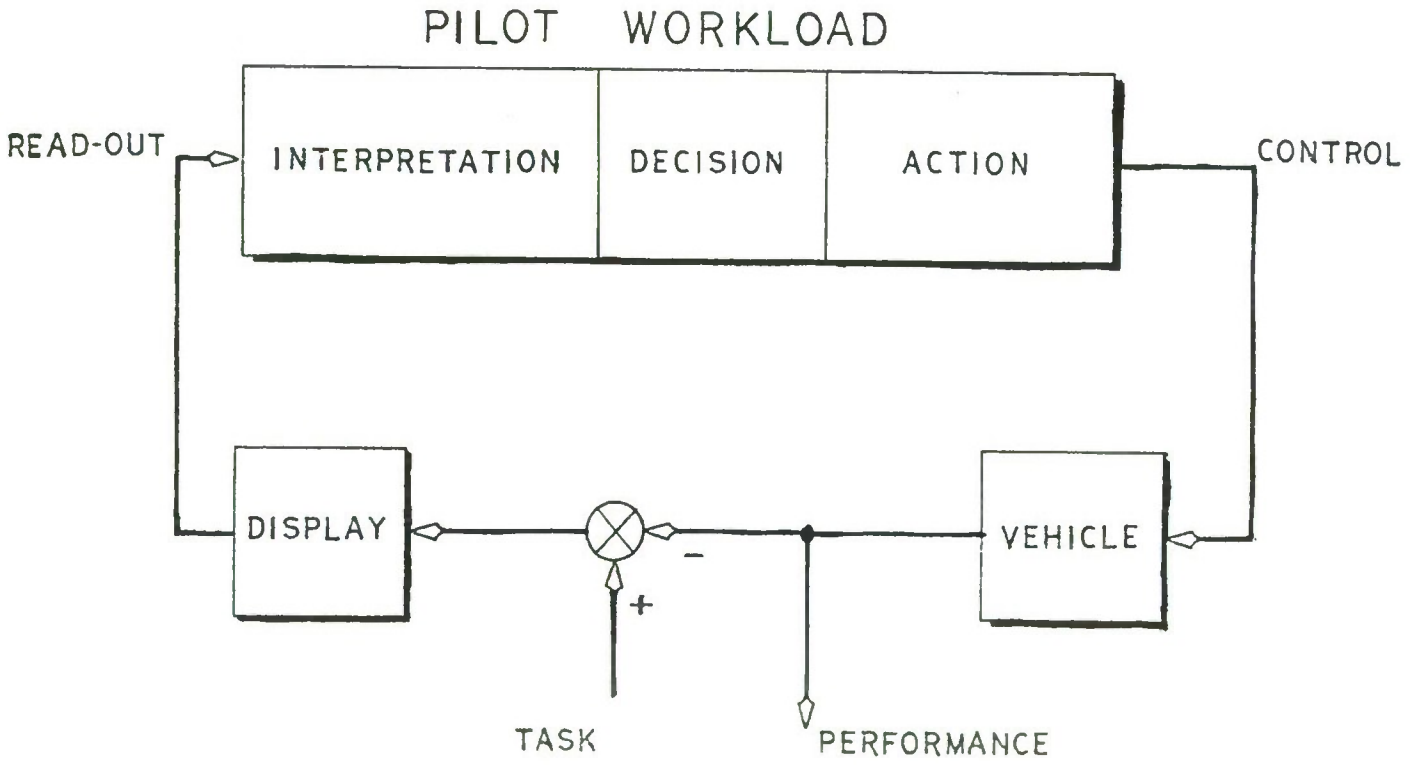


Fig. 1. Functional representation of the pilot's task.

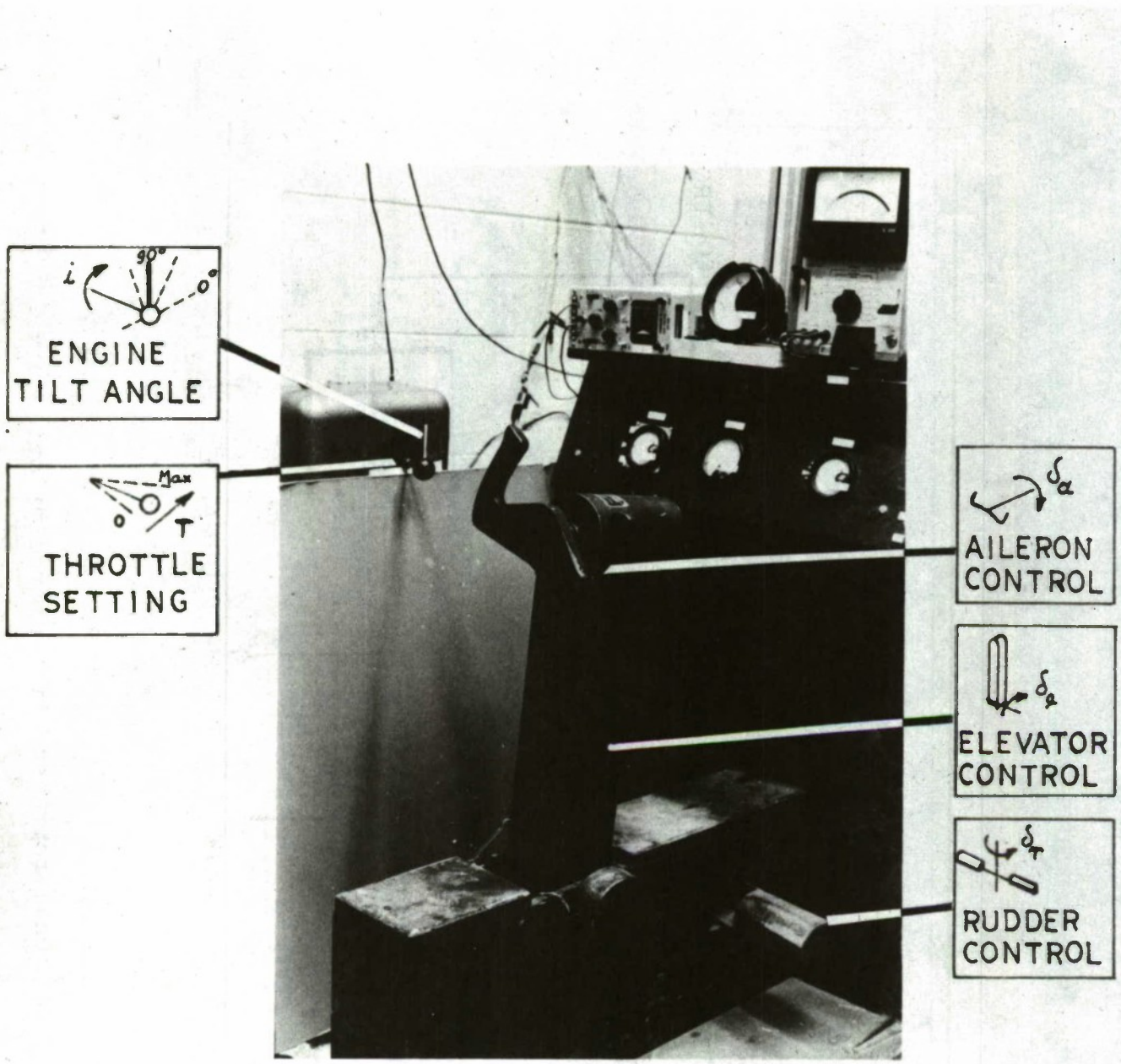


Fig. 2. The Controls in the Fixed Base Simulator.

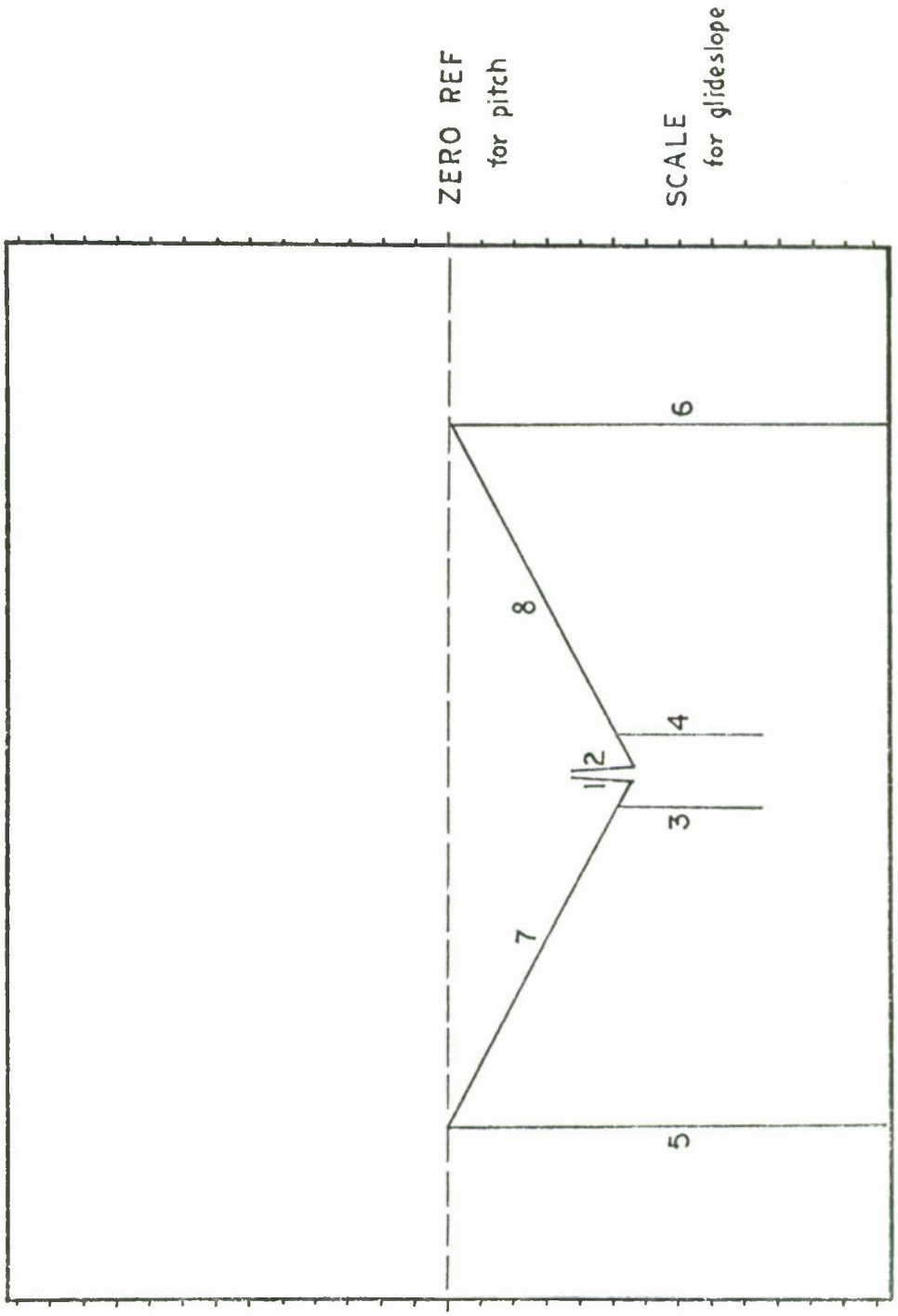


Fig. 3. The perspective display : scheme one.

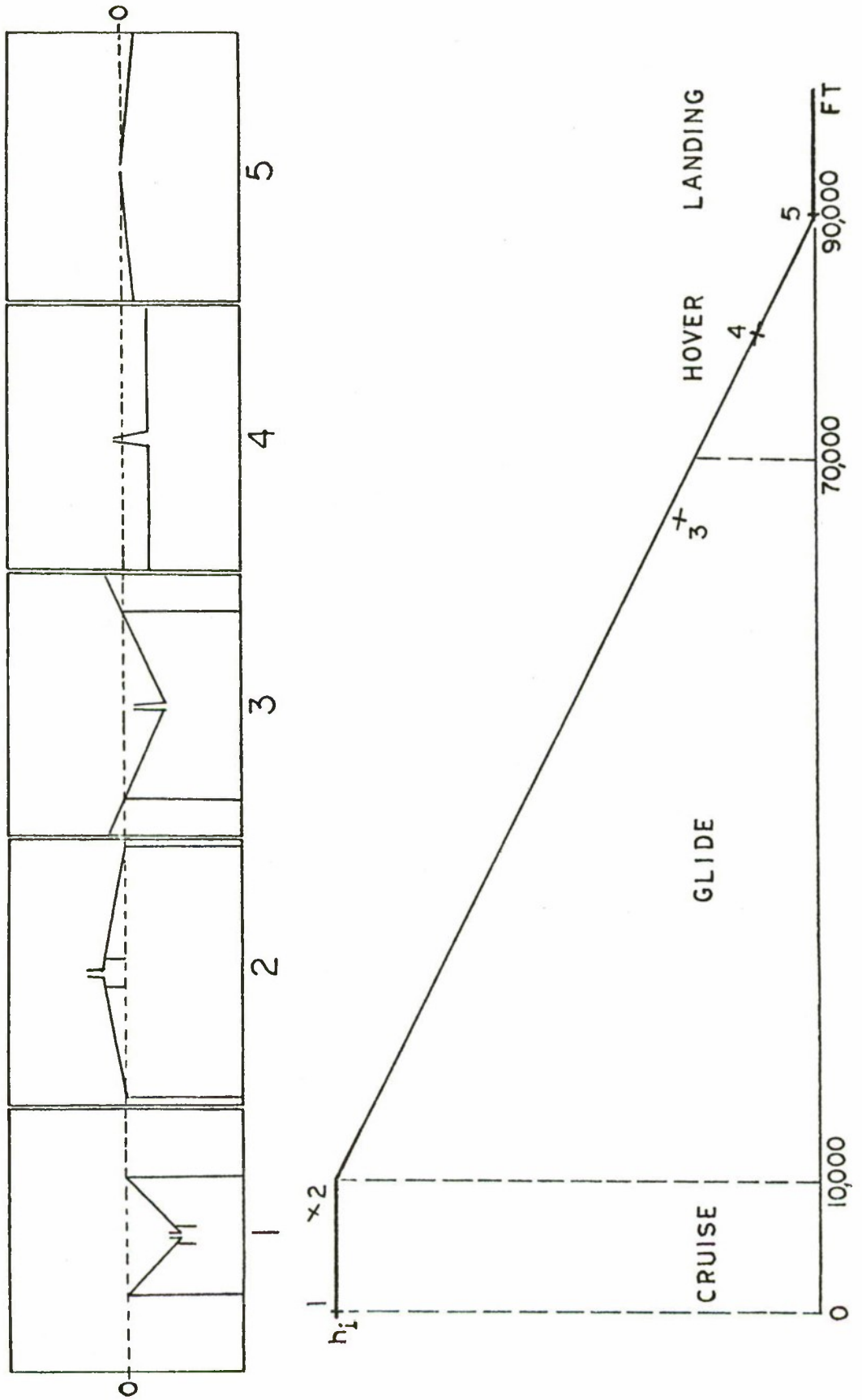


Fig. 4. The different views as seen from the indicated positions. (It is assumed that one is on the localizer beam.)

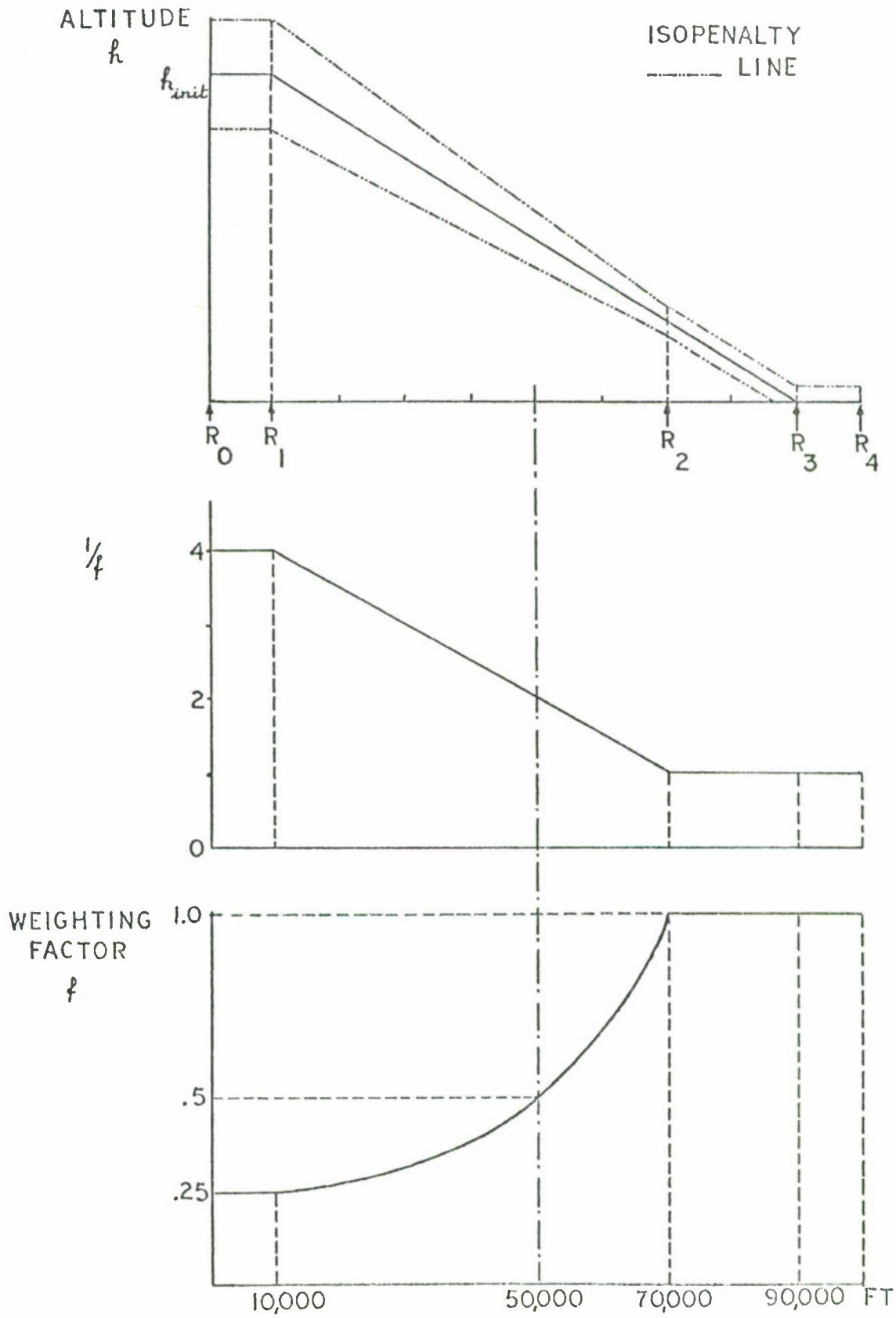


Fig. 5. Weighting factor for altitude deviation.

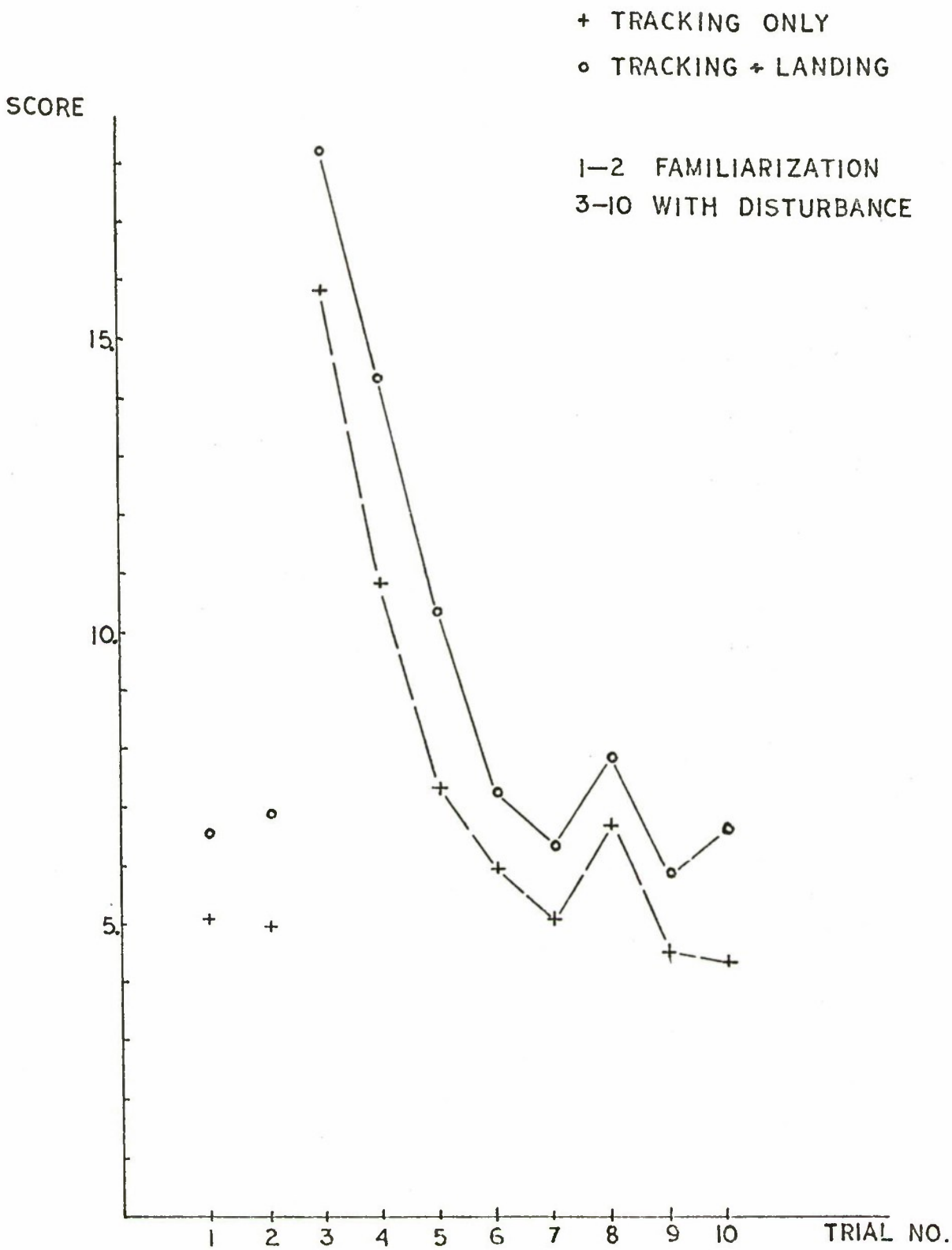


Fig. 6. Learning curve for an experienced pilot, using the conventional instruments.

Fig. 7. Range at touchdown in the case of the dial instruments.

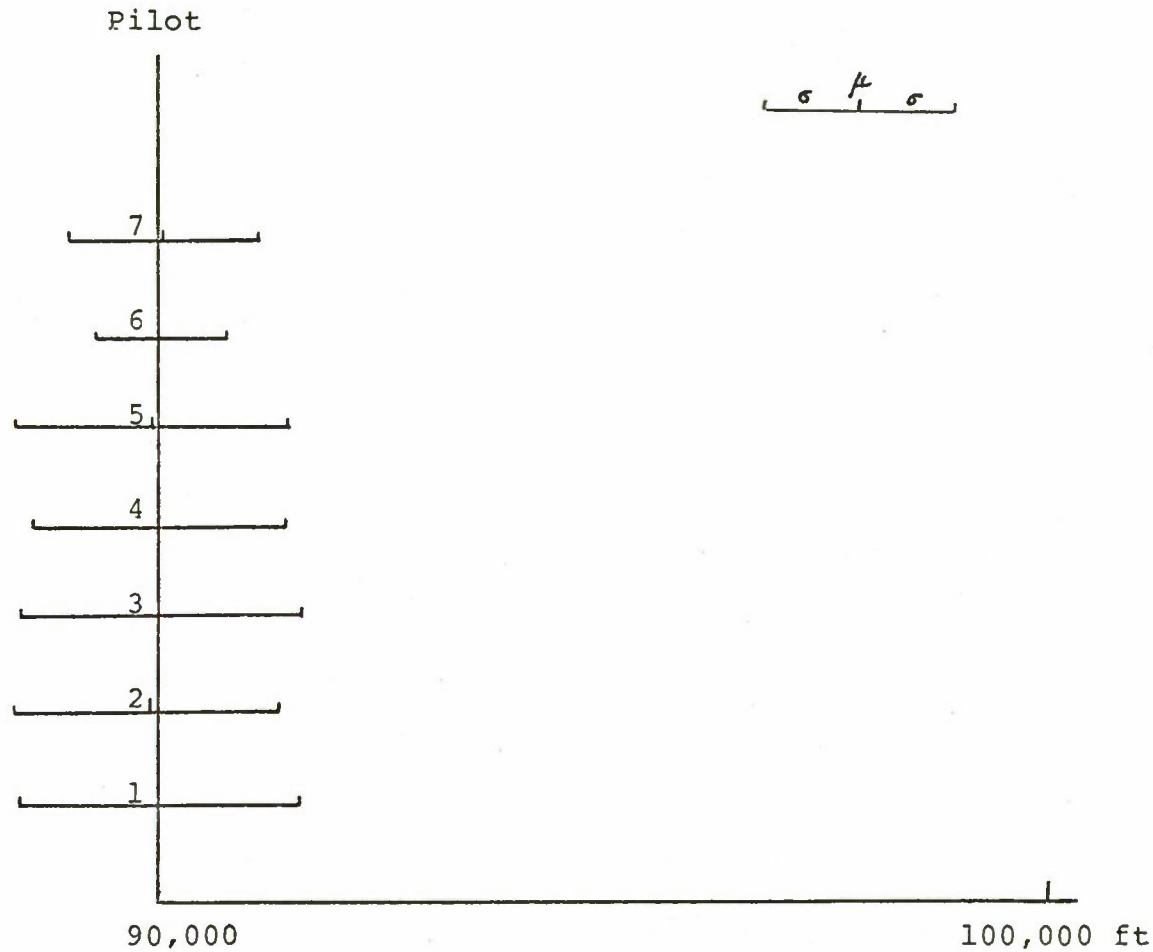


Fig. 8 . Range at touchdown in the case of the perspective display.

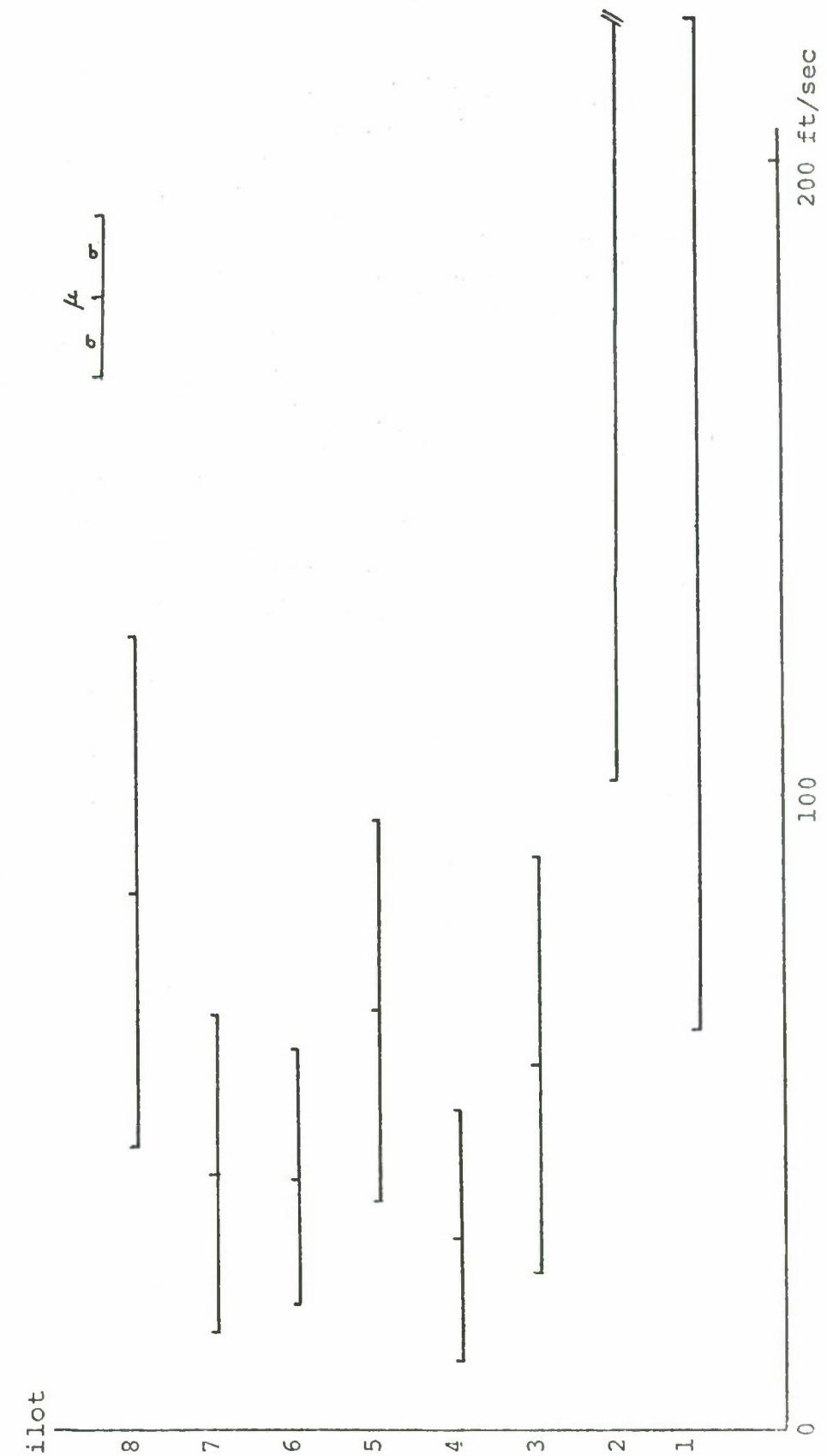


Fig. 9. Vertical touchdown velocity in the case of the dial instruments.

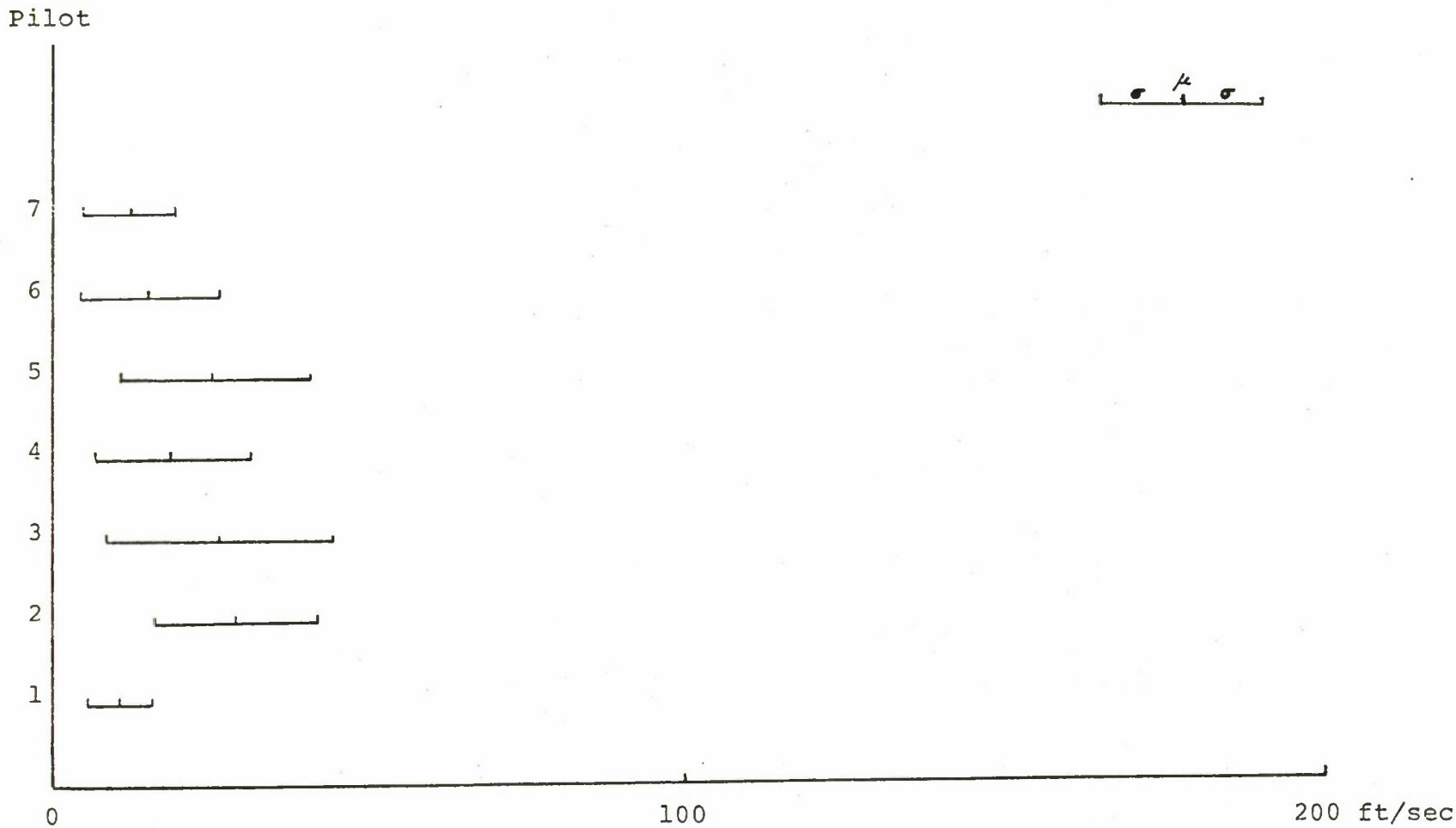


Fig. 10. Vertical touchdown velocity in the case of the perspective display.

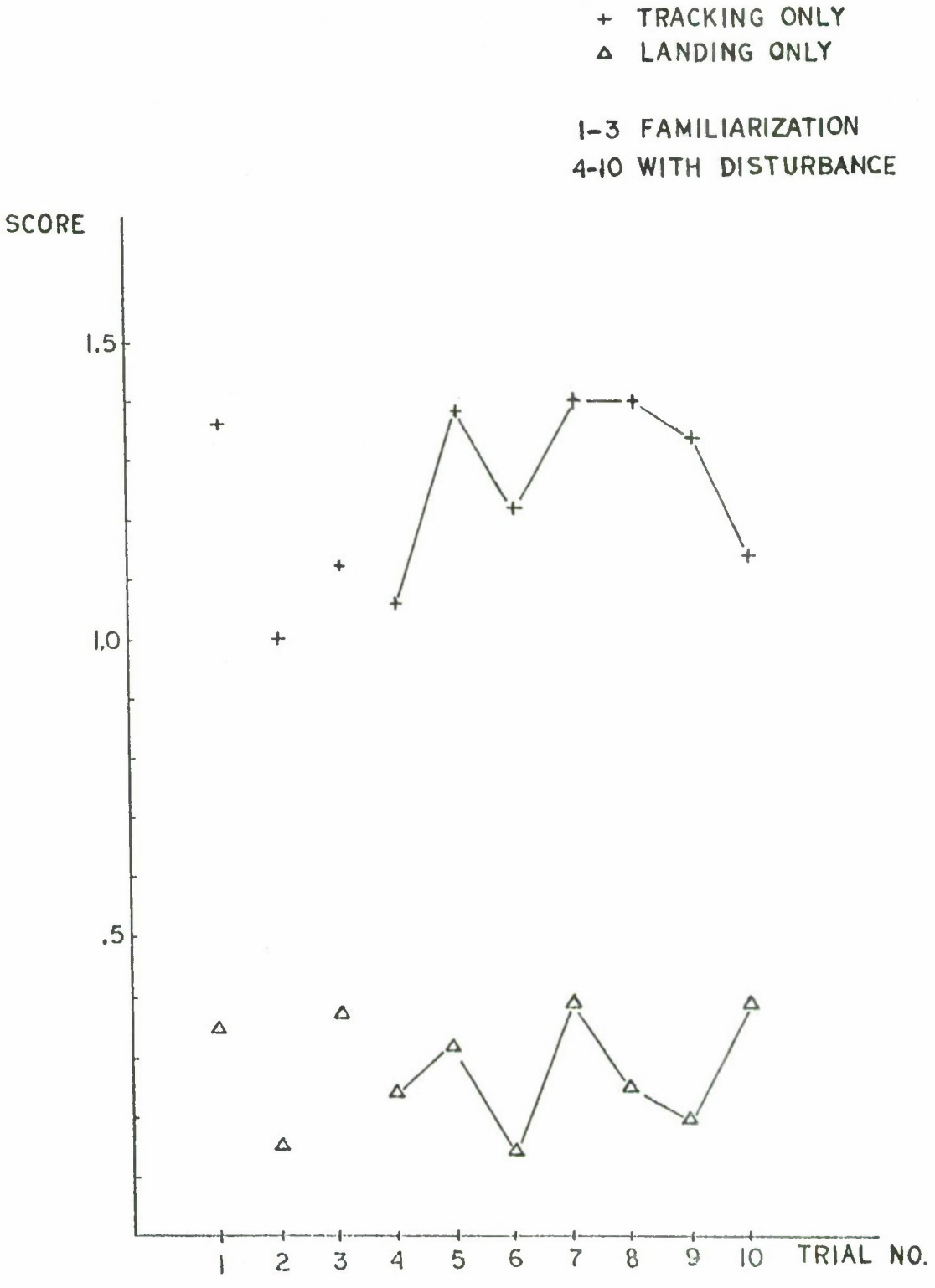


Fig. 11. Learning curve for an inexperienced pilot, using the perspective glideslope display.

AN INTEGRATED DISPLAY FOR TRAJECTORY CONTROL OF HELICOPTERS AND VTOL AIRCRAFT

Theodor A. Dukes

Princeton University
Instrumentation and Control Laboratory
Department of Aerospace and Mechanical Sciences
James Forrestal Campus
Princeton, New Jersey 08540

Presented at the
International Symposium on Man-Machine Systems
September 8-12, 1969, Cambridge, England

The work leading to and connected with the station keeping display and the simulator tests described in this paper have been supported by the U. S. Army Electronics Command under the Princeton Pennsylvania Army Avionics Research (PPAAR) Program, Contract No. DA28-043 AMC02412(E).

1 INTRODUCTION

Future generations of helicopters and V/STOL aircraft will have to satisfy requirements beyond their present capabilities in many respects. A family of significant requirements is related to the accurate control of the flight path under IFR conditions. The importance of presenting the pilot with the many elements of the necessary information in the most favorable way can hardly be overemphasized. Proper displays for steep descent landings, terrain avoidance, terrain following, collision avoidance and station keeping, have been recognized for many years as pivotal problems. At the end of the paper references are given in which reviews of past display work can be found. Some of them contain extensive bibliographies which may be consulted for detailed background information.

Flying helicopters and VTOL aircraft is more difficult than piloting conventional fixed wing aircraft for several reasons. Perhaps the most significant ones are the following: a) At and near zero translational airspeed (hovering in calm air) the dynamic behavior is essentially different and inherently unstable. b) The capability and need to accelerate and decelerate in flight between hover and cruise speeds requires a pilot to control a vehicle with rapidly varying dynamics. c) The in-flight change of configuration of a VTOL aircraft, the flaps of a STOL aircraft, a separate pusher engine of a compound helicopter, represent additional controls. The resulting non-uniqueness means that in a certain speed range the same translational flight condition can be achieved with various combinations of control settings.

These differences make precision flying under IFR conditions even more difficult as the translational accuracy requirements increase. With a larger number of essentially continuous positioning loop closures needed to perform a task, more information flow and more skill is required. In cruising flight only altitude has to be held with positioning accuracy. In a landing approach, vertical and lateral errors with respect to the nominal glide path center must be kept small. If a helicopter or VTOL aircraft is to be brought to hover over a point, predictive positioning in all three translational degrees of freedom is required. Such a problem, involving continuous control of the accelerations for positioning along three axes is usually not encountered in flying fixed wing aircraft, except for station keeping in formation. Precision hovering too, requires accurate position control and therefore continuous position loop closures in three degrees of freedom.

It has been recognized for years and has been shown conclusively [11] that for precision flying under IFR conditions integration of elements of displayed information is absolutely essential; a number of operational and experimental integrated aircraft displays are now in existence. Most of these displays belong in one of two classes. In one class, pilot input oriented display systems intend to arrange the presentation of the input variables closely related to the way they appear to the pilot under visual flight conditions. Flight director type displays, on the other hand, intend to tell the pilot what to do and are control oriented in a command sense.

The control motion oriented display concept to be described here combines several aspects of existing displays and some new ideas in a new format. It is based on a true presentation of the pilot input variables arranged in a single display, suggesting the control motions the pilot should make.

2 DISPLAY PRINCIPLES

The common denominator of all precision flying tasks is trajectory control which can include positioning along the desired flight path. The nominal trajectory can

be determined by radio beams, by the contour of the terrain (in terrain following), by an inertial system, or by another aircraft.

The integrated display concept presented in this paper has evolved from the general requirement of precision flight path control in all three translational degrees of freedom. It must be recognized that this general requirement is common to all flying tasks, only the shape and inclination of the desired flight path and the accuracy requirements are different in various tasks. The specific display to be described here is applied to helicopters. The concept and the basic features of the display have much wider applicability and can be implemented for compound helicopters and converting VTOL aircraft as well.

The position and the translational velocity components of any vehicle are controlled by applying proper acceleration components. The means of applying accelerations in the horizontal plane is tilting the resultant vector which balances weight and drag in equilibrium. Pitching forward provides the dominant part of longitudinal acceleration, and lateral accelerations with respect to a nominal flight path are achieved mostly by assuming a bank angle while the rudder control is used for coordination. The cyclic control stick motions of a helicopter with which the corresponding moments are applied, are indicated in the sketch on Figure 1. The direction of the control motion corresponds to the desired acceleration, but most of the acceleration occurs only when the attitude has changed. A sufficiently tight attitude loop closure can separate the dominant attitude and position response frequencies to the extent that the horizontal translational responses appear controlled by the pitch and bank angles.

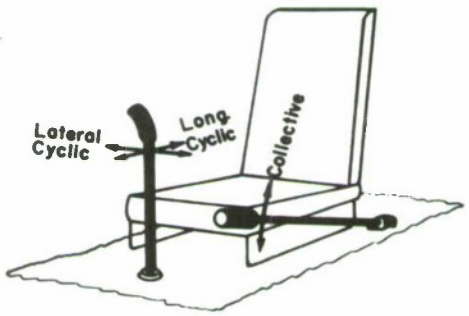


Figure 1. Helicopter Controls

The dominant control for vertical acceleration of a helicopter is the collective stick (Figure 1). This is essentially a power control, assuming that a governor or the pilot holds the rotor rpm nearly constant. Again, the direction of the control motion corresponds to the desired acceleration. The most difficult piloting tasks are those during which the collective stick cannot be simply set from time to time and then left alone for a while but must be used in a closed loop fashion. Such tasks are hovering, station keeping, the final phase of landing, and following a rough terrain.

For precision flying, loop closures involving these controls have to be performed nearly continuously with the intention to null the deviations from desired values of the respective variables. Only the deviations or errors are needed for these loop closures, the absolute values of the variables are of only secondary significance.

The integrated error display to be described in this paper is based on the following display principles.

2.1 Fully Integrated Multi-variable Information

This principle is aimed at reducing the scanning and fixation losses to a minimum. All the information necessary to perform continuous loop closures for controlling the aircraft in all three translational degrees of freedom should be available on a single display. Besides the error display, this includes all other information needed to perform a specific task, in particular information

for inner loop closures as well as augmentation of the basic display with additional symbolism on the environment. Aircraft nearby should be displayed if that information is available and ground information is necessary for landing or terrain avoidance.

2.2 True Deviation Display

This principle is aimed at minimizing the need for internal lead generation by the pilot and at establishing his confidence in the display by providing various cross checking possibilities within the display. Flight variables should be displayed in the form of deviations from desired values whenever possible. The pilot is to be presented with actual information. Rates are to be displayed explicitly.

The last two principles are aimed at a possible simplification of the pilot's internal signal processing.

2.3 Control Motion Oriented Arrangement

Elements of the integrated display are to be arranged so that their relative locations on the display and their motions in response to control inputs are easily and naturally associated with the respective locations and motions of the controls. In the case of conventional helicopter controls (Figure 1), this suggests a central area for the error presentation in the horizontal plane corresponding to the primary control of accelerations in this plane by means of the cyclic stick. The vertical error display should be to the left of the horizontal display. For application in other aircraft this arrangement or the controls might be modified to match the directions of control motions with the horizontal and vertical accelerations.

2.4 Systematic Arrangement of Derivative Information

The dynamic motion of symbols representing accelerations, error rates, and position errors should be in the same direction, leading each other respectively, so that their relative motions can be interpreted easily and naturally.

The rest of the paper describes an implementation of these design principles and some simulator test results. The paper is not concerned with the instrumentation and computation necessary to determine or approximate the information used in the display, the discussion here is confined to the display itself. As to the symbols used in the display, originality in symbolism was not a goal in applying the display principles; on the contrary, an attempt was made to use well established symbolism wherever possible.

3 THE BASIC DISPLAY

The configuration of the display as applied to the most difficult flying task, station keeping, is shown in Figure 2 and is used to describe the basic features of the display. It can be envisioned as an area for the translational information in front of the central area of a large attitude gyro. The position in the horizontal plane is displayed as the symbolic top view of a helicopter in a fixed coordinate system. The origin of this coordinate system is the "nominal point" on the desired flight path, and the x-axis (pointing up) represents the horizontal projection of the tangent to the desired flight path. The helicopter symbol in this error coordinate system also presents the deviation in heading from the x-axis. The position error along the vertical axis is displayed by a double bar at the left of the horizontal display area.

The size of the rotor symbol and the distance between the two altitude bars provide a sense of scaling of the display by representing the own rotor diameter to

scale and a fixed altitude difference of 50 feet, respectively. The rates of motion of these two symbols are presented explicitly by error rate vectors originating in the centers of the symbols. These vectors "predict" the motion of the helicopter and altitude error symbols. The vector tips are emphasized by means of small circles. Since the motions of the vector tip symbols on the display present the sum of position and rate, these vector tips are equivalent to quickened position displays. The explicit presentation of the error rate provides instantaneous rate information and eliminates the need to derive this information from the motion of the symbol. The fact that not only the resultant is shown, but position and rate are separately identifiable provides the pilot with true situation displays and lets him make decisions about how he wishes to make his corrections.

Information related to the accelerations in all three translational degrees of freedom is also displayed. The acceleration along the longitudinal axis is closely related to the pitch attitude of the helicopter. The horizon symbol and the "sky" are shown on both sides of the central display area described so far. The integration of the attitude gyro into the display is extremely significant since this makes it unnecessary for the pilot to refer to another instrument for the information which he needs most for his inner loop closures, his confidence, and his safety. The two pairs of bars which roll together with the horizon are reference symbols for the pitch attitude and are ajustable by the pilot or could be used as flight director inputs. The lateral acceleration with respect to a nominal flight path can be represented by a turn needle symbol which is seen at the bottom of the central display area. A ball indicator is also added for completeness and represents lateral acceleration at and near hover. The second derivative of the vertical displacement is basically the thrust variation, and as a measure of this a small symbol showing the deviation of the collective stick position from a set reference has been included at the left of the altitude information. An alternative for this symbol is to represent torque rather than collective position.

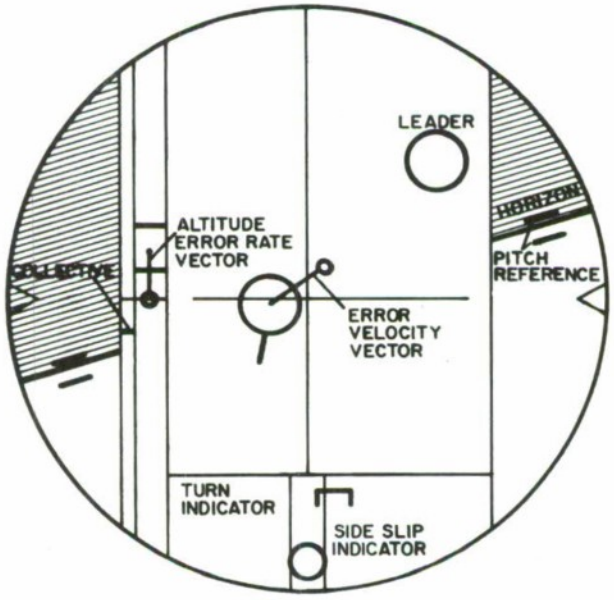


Figure 2. Integrated Display, Station Keeping Mode.

It should be noted that in all three degrees of freedom the symbols representing the derivatives move in the proper sense with respect to each other when a control action is initiated.

The symbolism shown in Figure 2 was implemented on a cathode ray tube. In the course of initial experiments it became quite obvious that the continuous observation of a large number of symbols on a single display can become rather tiresome after a while. A multi-color presentation* of the display was developed

---

* A multi-color cathode ray tube for initial experiments was provided through the courtesy of Sylvania Electric Products, Inc.

and was found vastly superior. Three distinct shades of color are used for the separation of families of symbols. A low intensity orange color is used for all reference and border symbols including the two pairs of pitch reference bars. Most of the symbols presenting variable information are bright yellow-green. Based on theoretical and experimental evidence, four symbols of greatest significance were singled out to catch the pilot's attention: the two vector tip circles and the two horizon segments. These are presented in red. The value of using the red color to emphasize the vector tips and the horizon was explored experimentally. It became quite clear that the three-color version results in a considerably better display than a two-color presentation.

The pilot has the following controls over the station keeping display:

Mode selection
Range selection
Pitch reference adjustment
Bias of collective reference
Intensity

4 DISPLAY MODES

The integrated symbolic display described in the preceding section contains all the information necessary for the continuous loop closures to fly the aircraft. The result of applying the stated display principles is the combination of vertical and horizontal information into a single display in a systematic arrangement, related to the pilot's controls both geometrically and dynamically.

The basic principles, the arrangement, and most of the symbolism should be kept invariant if possible regardless of what the task and the corresponding display mode should be. The advantage of such commonality is clear: pilots could be trained in IFR flight path control and use basically the same internal processing between their input (the display) and their output (the control motions) throughout IFR flying without referring to a variety of different displays and display principles. In Reference [12] some ideas are described, how the same basic display principles can be applied to different flying tasks and how the basic display configuration might be augmented to satisfy special requirements in each case. Such special requirements are mostly for environmental information both for safety and as a basis for decision making if and when modification of the desired flight path becomes necessary.

For station keeping, e.g., the addition of a leader symbol provides the required sense of proximity. Any other aircraft within the boundaries of the displayed airspace must also be shown.

Of the other display modes proposed in References [12] the landing mode is described in some detail below. A consistent application of the stated principles to approach and landing results in a break with the conventional basic ILS display. In order to keep the same close coordination between display and control motions under all flying conditions, the fundamental arrangement of horizontal plane and vertical axis should be maintained. The input obtained from the localizer signal, possibly with course softening, determines the lateral error. Longitudinally the helicopter symbol does not leave the reference axis since there is no need for continuous position loop closure along the glide path. The longitudinal error rate vector component is to be determined by the deviation from a set airspeed. The deviation from the glideslope is the input to the

vertical error display.

It is proposed that for landing the basic integrated display should be augmented in two ways. In the vertical error display strip a ground proximity line (in red) should be moving up from the bottom for the last 200 feet of altitude above the touch down point, as an "absolute" altitude indicator (see Figure 3). The other added symbolism, shown in the top half of the central display area in Figure 3, is to be interpreted as follows. Imagine the view of the touch down line during approach as seen from a point fixed to, but above, the helicopter. During descent, this line is seen to come closer to the nominal point, and at the same time it grows compared with the top view of the helicopter itself. Its endpoints would describe two loci on the display as certain nonlinear curves if the helicopter would fly exactly on the glideslope. If these loci are shown on the display they can be used as reference lines for displaying the glideslope error. If the helicopter is below the glideslope the view of the touch down line extends beyond the reference lines; if the helicopter is above the glideslope the end points of the touch down line do not reach the reference lines. On the proposed relatively simple version of the landing display the reference lines are straight instead of curved without any loss of actual information. This augmentation of the display combines an indication of the distance to touch down with a display of vertical error.

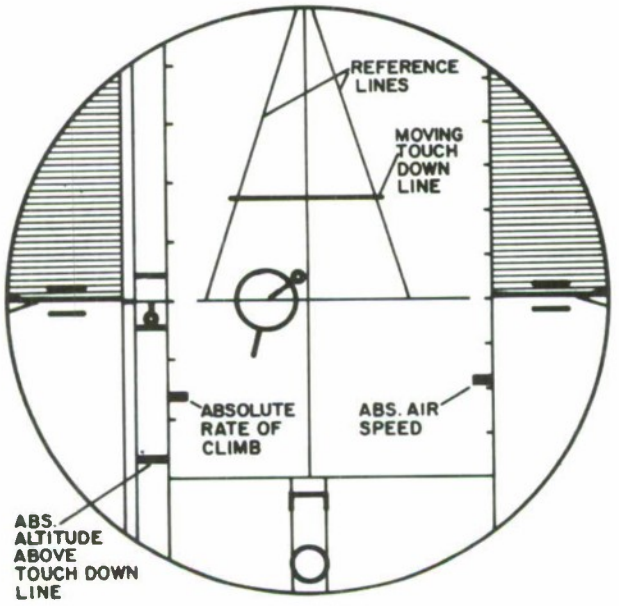


Figure 3. Integrated Display, Proposed Landing Mode.

If adequate instrumentation is available, transition to hover could be commanded through the reference of the longitudinal component of the error rate vector, augmented by a pitch command, if necessary.

In order to avoid the need for scanning other instruments provision has been made to add two absolute scales to the display as indicated in Figure 3. Each division on the absolute rate of climb scale represents 500 feet/minute, each division on the absolute airspeed scale represents 25 mph. Whether the addition of this information is worth the increased density of symbols has yet to be explored.

In a cruise and terrain avoidance mode, the moving touch down line and its reference lines could be replaced by lines indicating the ground contour (see Reference [12]). It has been proposed by an experienced pilot that a radar display could also be superimposed on the horizontal display area. In the cruise mode, the helicopter symbol would not move away from the origin of the reference coordinate system in either direction. The orientation of the tail shows the difference between actual and desired heading. The longitudinal component of the rate vector is proportional to the difference between actual and desired airspeed. The lateral component of the rate vector could be used to display a measure of "course error" since this indicates the rate with which the helicopter is moving away from a desired flight path.

The following controls over the display must be available to the pilot in addition to those mentioned in the preceding section:

Course reference setting
Altitude reference setting
Airspeed reference setting

The same basic display configuration can be maintained in all display modes which should have a beneficial effect on instrument flight training.

5 SIMULATOR TESTS

5.1 Test Description

The present format of the described integrated error display has evolved through many hours of simulator testing. In these tests a fixed-base cockpit was used in conjunction with an analog computer on which the helicopter equations were set up and the flight variables were transformed into the display coordinate system. The helicopter equations were those of a UH1 flying at 96 miles per hour augmented with some pitch and roll damping, as used in similar explorations of other displays and described in References [8 & 9]. The dynamics of the helicopter are illustrated in Figure 4 by the three responses to a longitudinal and a lateral cyclic pulse and a collective step input, respectively. The strong dynamic coupling between the horizontal longitudinal and the vertical responses x_e and z_e indicates the difficulty in holding position in all three translational degrees of freedom.

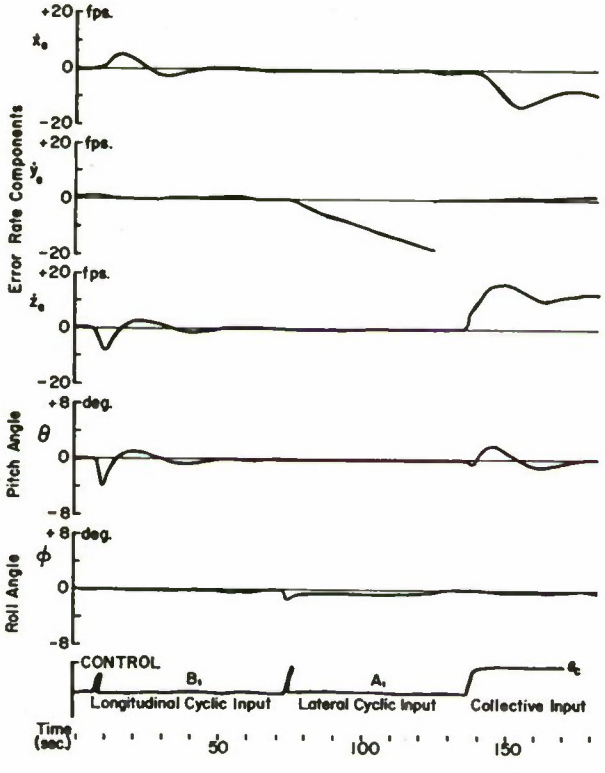


Figure 4. Responses of Helicopter Simulation Model (UH1, 150 ft/sec, with Added Pitch and Roll Damping)

In most of the simulator tests the task was station keeping with respect to a leader, using a 5-inch diameter display. The nominal point was three rotor diameters (150 feet), to the left rear of the leader. The requirements were somewhat simplified compared with actual "rigid" formation flight, inasmuch as the leader's flight path was assumed to be known and disturbance free. The nominal flight path of the follower was taken parallel to the leader's flight path, eliminating the initial "swing-out" which occurs when a rigid formation is referenced to the leader's heading. Instrumentation noise was also neglected.

Gust disturbances were not simulated for the following reasons. In a tight formation it is common practice in visual flight to "ride out" most of the displacement effects of gusts since, because of the close proximity, the gust inputs to leader and follower are very similar, and only the relative displacement matters in the station keeping task. Both leader and follower independently counteract the gust effects on pitch and roll attitude and on heading. As a result, only the effects of the high frequency end of the gust spectrum should be considered. On the other hand, in actual flight the high frequency accelerations can be

sensed by the pilot directly and this can provide valuable lead information. In a fixed base simulator the pilot is faced with a quite unrealistic situation if the display shows him effects of accelerations which he does not feel. This lack of realism, together with the observation that after the application of an input the pilot induced disturbances are enough to keep him active all the time, was considered sufficient reason for omitting gust disturbances in the initial evaluation of significant display parameters.

5.2 Test Results

Several qualitative features were established based on pilots' comments and observations. It was found that the relative intensity of symbols can be almost as significant as color coding. For example, high intensity needle and ball symbols distract the pilot's attention despite their peripheral location. As a result of other comments, a distinct separation of the control display area and the "gyro" has been emphasized.

One of the most important questions was: how significant is the presentation of the artificial horizon on the same display? It is conceivable that the pilot can fly the vector tip alone, the displacement errors and the attitudes being only one integration and one differentiation away, respectively, and these can be generated by the pilot. It actually has been demonstrated consistently by several pilots that with the horizon blanked, the station keeping task can be performed as accurately as with the horizon. However, without the horizon, large amplitude attitude excursions occurred in the process. Such large excursions would never be tolerated in flight. These excursions were reduced when a separate attitude gyro was used, at the expense of lower positioning accuracy and higher work load. The integration of the gyro and the station keeping display area was recognized as an asset by all test subjects. In order to enhance "seeing without looking", the "sky" is shaded by means of a scanning beam so that motion of the horizon is recognized easier through parafoveal vision. This feature has not yet been incorporated in the tests.

There are many aspects of less significance which can hardly be explored conclusively. But there are two important questions which demand quantitative answers: what are the best error display ranges and what should the relative vector gains be? The choice of range in the horizontal plane is limited on the low side by the requirement that the leader must be displayed within the range and that the minimum practical safe nominal distance in actual IFR flight is expected to be not much less than three rotor diameters. With increasing range, the poorer resolution can be expected to result in less accurate performance.

A vector "gain" K means that the length of a vector corresponding to an error rate of x feet/sec is K.x feet long on the display scale. It was clearly established during preliminary tests that the pilots dwell most of the time on the vector tips. Therefore an attempt must be made to optimize the vector gains. An optimum clearly should exist since with a too low gain the vector would hardly be distinguishable, and with a too high gain it would move about too "nervously", especially in the presence of unavoidable gust disturbances and instrumentation noise.

A limited test series was conducted in order to explore quantitatively the effects of scales and vector gains in the horizontal plane. The test subjects were six U.S. Army pilots with helicopter flying experiences ranging from 400 to 5,000 hours. Based on preliminary simulations, two different scales and four different vector gains were chosen. The scales were 125 and 250 feet, referring to the

distance between the nominal point and the right-hand-side boundary of the horizontal display area. The vector gains were 3.0, 5.2, 9.0 and 15.0, approximating a geometric series with the factor $\sqrt{3}$. Each pilot performed twice with each combination so that the results of 12 runs could be averaged for each test point. A run was made up of five distinct segments, beginning and finishing with straight-and-level flight, with one left or right turn, and one climb or descent segment, separated by a third straight-and-level segment. The duration of a run was about 3 to 4 minutes. In Figure 5 two test runs illustrate the differences in pilot performance: Figure 5a represents a very good performance and Figure 5b shows a poorer than average performance with the same test variables. These runs also demonstrate that the lengths of the individual segments are not sufficient to establish steady state characteristics; the segments are dominated by the translational transients. The emphasis on the transients was deliberate because they represent the most crucial tests of performance. Another significant aspect of the display, which is connected with the question of fatigue during a long mission in mostly straight-and-level flight, requires a very long test series in actual flight or at least in a moving base simulator.

Singling out the most significant segments, the data of the turn phase and of the climb/descent phase of each run were examined first, and some results of this data reduction are presented here. Both peak deviations and averages were derived from the data.

Figure 6 shows the distributions of maximum deviations along the x, y, and z display axes, taken from all test runs. (The combination of the 250 ft. scale and the vector gain of 3 was omitted after initial runs had shown that this is a clearly inferior combination.) The distribution patterns appear distinct enough to indicate that the 12 runs per test point are adequate. In evaluating the results it should be kept in mind that the x and z errors in a turn and the y error during a

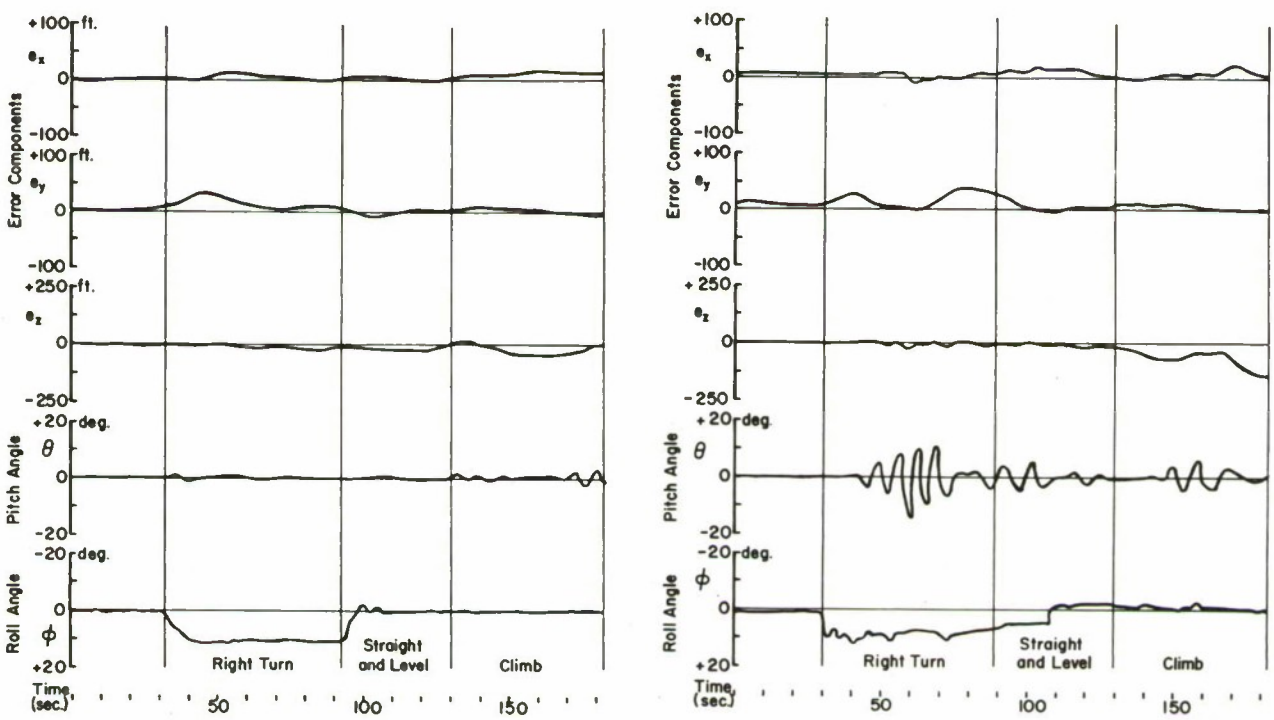
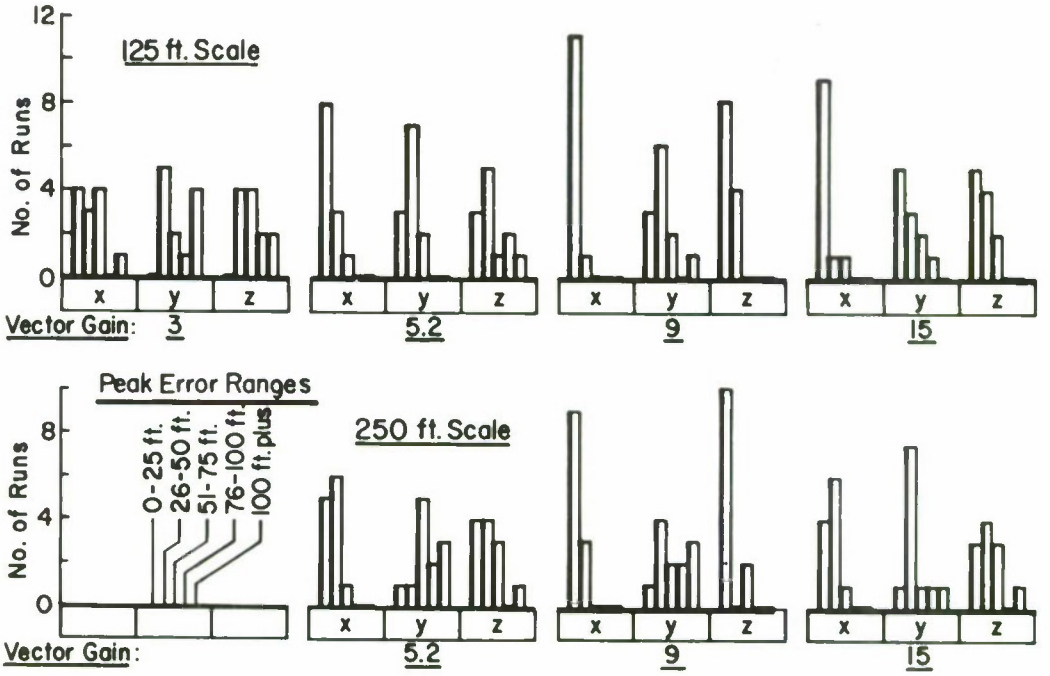
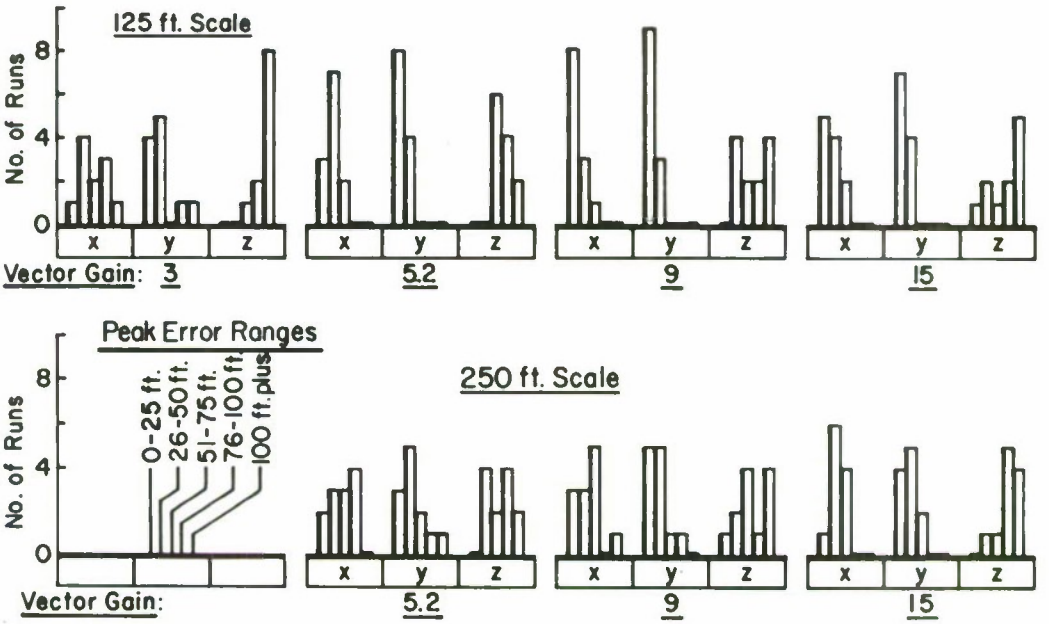


Figure 5. Typical Records of Station Keeping Simulation.



a.
Turn



b.
Climb/Descent

Figure 6. Distributions of Maximum Error Components During Maneuvers in Station Keeping.

climb or descent correspond essentially to undisturbed loop performances, since the maneuver transients occur predominantly in the y and x, z axes, respectively.

Based on the maximum errors during maneuver transients as shown in Figure 6, using the 125 ft. scale results in distinctly better performance at any vector gain. With this scale, the vector gain of 9 seems to have an edge over the two adjacent gains, 5.2 and 15. With this best gain the errors in the essentially undisturbed degrees of freedom never exceeded 50 feet.

The rms values of the x, y, z error components and of the range R in each maneuver segment also have been calculated and their averages are shown in Figure 7. The patterns of these averages are consistent with those of the peak error distributions in Figure 6.

Considering the significance of large errors in the station keeping task and the relatively poor performance with all vector gains when the 250 ft. scale was used in the climb/descent maneuver, it can be concluded that the 125 ft. scale is better for overall station keeping performance. This conclusion is in agreement with the visual flying preference to fly as close to the leader as safety permits.

The adequacy of the learning period of approximately three hours was checked by averaging separately the first and second times each test run was performed; no significant differences were found. Similarly, averaging separately and then comparing the data obtained with young pilots and with very experienced pilots, also showed no significant differences.

It is interesting to note the effect of the test variables (scale and vector gain) on the z-axis performance since the vertical error scale and error rate vector gain were held invariant throughout the test series. This indicates that the pilot's attention pattern is influenced by significant variation of any display parameter, affecting all closed loop performances.

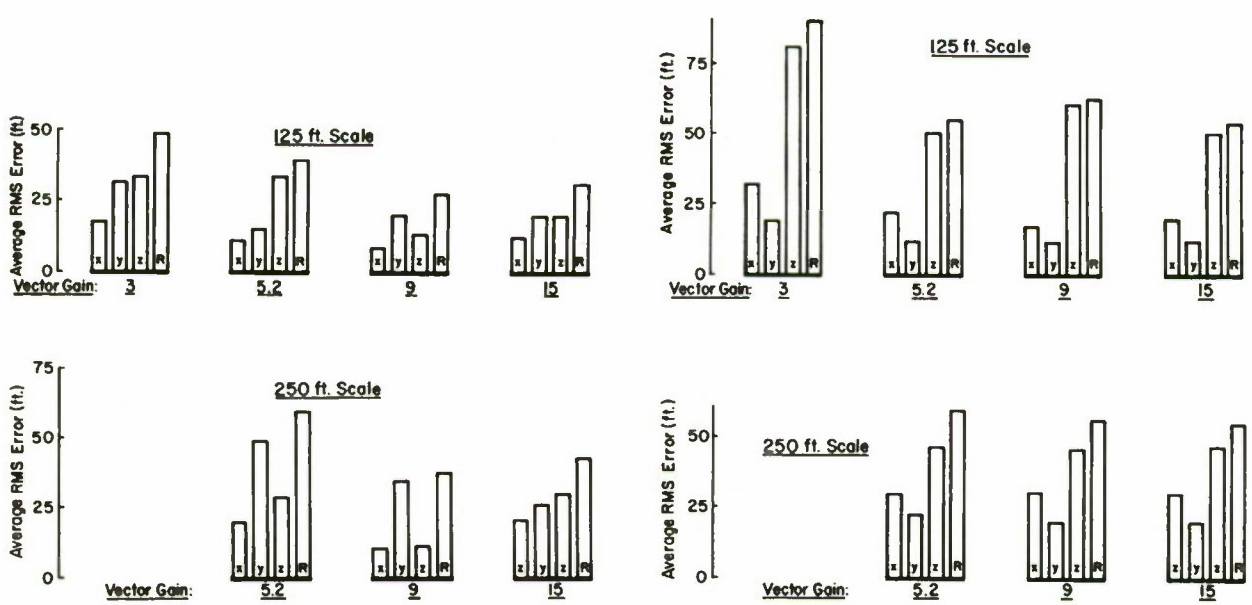


Figure 7. Average RMS Error Components and Range Errors During Maneuvers in Station Keeping.

The relatively large z errors, especially in the climb/descent maneuver, can be partly attributed to the fact that in station keeping the pilots are not so much concerned about vertical deviations as they are about errors in the horizontal plane. Also, no attempt has been made in this test series to optimize scale and vector gain in the vertical error display region.

Additional clues available in actual flight might aid the pilot in his inner loop closures and could result in some improvement along one or more axes. For further explorations an experimental airborne version of the display is being built to be tested in a flying simulator.

6 CONCLUSIONS

Integration of all necessary flight information for precision trajectory control of helicopters into a single display has been achieved successfully, adhering to a set of simple display principles. A basic invariant display configuration can be augmented to apply to any precision flying task and is adaptable to other VTOL aircraft.

Fixed-base simulator experience with a 5-inch display was obtained in the station keeping task which requires accurate position control in all three translational degrees of freedom. The test variables were the scale and the error rate vector gain in the horizontal plane, with the vertical display parameters held invariant. The performance was conclusively better using the smaller (125 ft.) scale. This is a lower limit for station keeping but an even smaller scale might be beneficial, e.g. for precision hovering, instrumentation noise permitting. Among the error rate vector gains tested (3, 5.2, 9, 15) the gain of 9 appears best,but not as a very sharp optimum. Actual flight testing, in the presence of instrumentation noise and gust disturbances, may easily shift the optimum to the next lower gain. Further improvements in the vertical display parameters are necessary and will be explored in the next test series.

REFERENCES

1. Anast, James, Advanced Cockpit Instrumentation, North Atlantic Treaty Organization, AGARD Report No. 235, May 1959.

2. Brown, E. L., The U.S.A.F. Vertical Instrument Program, North Atlantic Treaty Organization, AGARD Report No. 236, May 1959.

3. Burke, J. E., Huchingson, R. D., Koppa, R. J. and Sewell, H. E., Jr., First and Second Simulator Evaluations of Advanced Integrated Display and Control Systems, NASA CR 762, June 1967.

4. Cole, G., Bender, M., Shoquist, R., Santella, R., Lovinger, D. and Lindquist, O. H., Study of Pilot-Controller Integration for Emergency Conditions, Air Force Flight Dynamics Laboratory, Research and Technology Division, RTD TDR 63-4092, 1963.

5. Kelley, Charles R., Manual and Automatic Control, John Wiley & Sons Inc., New York, 1968.

6. Kingston, William W., Visual Display Techniques in Vehicle Stabilization and Control, In National Symposium on Information Display, 6th, New York, Sept. 29-30, 1965.

7. Koppa, R. J., A Survey of Current Display and Control Concepts, Ling-Temco-Vought, Inc., Astronautics Division, Report No. 00.308 (NASA Contract NASw-611), March 1964.

8. Nicholson, R. M., Wolf, J. D., Clifford, R. R. et.al. (Honeywell, Inc), Display Requirements for Helicopter IFR Formation Flight, JANAIR, ONR NR2-3-054, 1968.

9. Nicholson, R. M., et.al. (Honeywell, Inc.), Experimental Evaluation of a Formation Flight System, JANAIR, ONR TR-6804002, 1968.

10. Winblade, Roger L., Current Research on Advanced Cockpit Display Systems, North Atlantic Treaty Organization, AGARD Report No. 491, Oct. 1964.

11. McRuer, Duane, Jex, Henry R., Clement, Warren F., and Graham, Dunstan, A Systems Analysis Theory For Displays in Manual Control, JANAIR Contract No. N00014-66-C-0072, NR 213-044 (Systems Technology, Inc. Technical Report 163-1), October 1967, Revised June 1968.

12. Dukes, Theodor A., An Integrated Display Concept for Helicopters and VTOL Aircraft, American Helicopter Society Preprint # 314, 25th Annual National Forum Proceedings, Washington, D. C., May 1969.

THE STATISTICAL PHASE PLANE DISPLAY: A NEW METHOD FOR MONITORING TIME SERIES

by

JOSEPH G. WOHL

INFOTON INCORPORATED
Second Avenue
Burlington, Massachusetts 01803

ABSTRACT

The objective of the work to be reported on was to design, demonstrate and evaluate new methods for transforming and displaying time series information from a single data source. The time series themselves could represent the time histories of voltages, temperatures, and pressures in engineering applications; or respiration rate, heart rate, and blood pressure in medical applications. The goal of the effort was to develop a display format, suitable for CRT presentation, which would enable an individual to monitor a normally stationary (in the stochastic sense) data stream over a long period of time and to detect departures from stationarity (for example, shifts or trends in the mean value) with a minimum of human attention and computer involvement.

Several carefully selected display formats and data transformation techniques were developed and demonstrated in order to determine their efficacy in meeting the basic requirements of data compression, ease of interpretation, flexibility of use, and minimum operator, computer, and programming requirements. The interpretability of each display technique was judged primarily on the basis of the delay in detecting the shift in the mean of a computer-generated stream of simulated data with specified stochastic properties.

Of the techniques evaluated, the Statistical Phase Plane was judged to be superior to all others. This is a plot of the exponentially smoothed values of the first and second mements of the time series, plotted one against the other and displayed in real time.

Introduction

There exist a large number of applications in engineering, medicine and quality control in which many quite stable parameters must be monitored over long periods to check on whether they have in some sense become unstable. Many such continuous monitoring tasks are now being handled by computers, programmed to compare each monitored parameter against prestored go/no-go limits. The comparisons occur once every computer cycle. However, the computer takes no action unless and until a parameter exceeds its prestored limits, when the occurrence is brought to the attention of cognizant personnel. But such go/no-go information cannot suffice for a professional nurse, engineer or production controller; they must have additional information if they are to fulfill their responsibilities.

With the advent of computers and CRT terminals, it has become feasible to display a number of parameters for monitoring on a single CRT display. The present report is concerned with the general problem of data transformation and display formatting to enhance detection and interpretation of "within-limits" disturbances to otherwise stable data. A single parameter is assumed.

Such disturbances might be, for example, linear and nonlinear trends, sudden shifts, impulses, and periodic cyclical variations. For normally stable parameters, the problem is compounded by the phenomenon of human vigilance decrement [1] typically experienced in long-term monitoring of stable parameters: an observer observes no change; no change occurs; after a while he comes to expect no change; and when a change does occur, he does not notice it unless it is a drastic one.

Unfortunately, actual data "streams" from real systems were not available, so that it was necessary to simulate various data stream characteristics for demonstration and evaluation purposes.

Detailed results of the analysis and experimental work were presented in earlier reports [2,6,7], in which a number of different techniques were evaluated. The present paper summarizes the results of the earlier work and presents some new results.

Approach

The approach used was to first characterize the basic types of disturbances to a data stream, and to evaluate the utility of various data transformations and display formats for assisting observers in detecting the disturbances under simulated test conditions.

A simple yet precise way of characterizing a stable data stream is in terms of its stationarity. If the statistical properties of a portion of a data stream are identical to those of a later portion, we say that the data stream is stationary.[3] The technical problem may then be expressed as the development of methods of either testing directly for, or enhancing the "visibility" of, departures from stationarity. We assumed the general case of a data stream being sampled at a relatively low rate (e.g., once per second), and subject to random, small-amplitude, gaussian-distributed fluctuations due either to superposition of random noise or to underlying physical phenomena. Thus, small disturbances or departures from stationarity would tend to be masked by these random fluctuations, and the proposed techniques would be put to a stringent test.

Analysis

As noted earlier, departures from stationarity in a data stream may take several forms. These are:

Random, impulsive "noise"

A sudden shift from one stable state to another (e.g., "step" change in mean).

A gradual, continuing change in one or more statistical characteristics (e.g., monotonic trend in mean).

Oscillatory shifts.

Human ability to detect such changes using various display formats was evaluated in an earlier report.[1] These formats included meters, pen recorders, and alphanumeric CRT displays. The results indicated that the time plot format, exemplified by the typical pen recording trace, was by far the "richest" of the formats in general use, in that it provided the most clues to data interpretation.

The advent of centralized data monitoring via CRT display, together with increasing numbers of parameters to be monitored in many important applications and the decreasing availability of trained personnel, suggested that a number of hitherto little used methods of transforming and displaying data streams be evaluated for their applicability to computer-driven CRT displays for monitoring. In addition, it was considered possible, through application of appropriate mathematical operations, to

program computers for more sophisticated detection of these disturbances. Both approaches required mathematical transformation of the data stream, in the one instance to enable the computer to perform automated tests for the various types of non-stationarity and in the other instance to enhance the visibility of these non-stationary disturbances for trained personnel viewing CRT displays.

The cathode ray tube, when used in a tabular alphanumeric format, takes on all of the advantages and disadvantages of what might be called a multiple-digital readout display. Almc t everyone is familiar with this type of display, which has grown in use since the advent of digital computers. The digital readout indicates only the instantaneous state or value (to a selected number of decimal places of precision) of the measure being taken. In terms of legibility and precision, it has the advantage of being the most accurate form of display available. A second advantage of this type of data coding, of course, is that many such indications can be compressed into a high-density CRT display in a page-type format for centralized monitoring purposes.

However, introduction of the page-formatted CRT (see Figure 1 for a classic example) to assist in monitoring has posed a difficult problem. Whereas trained personnel formerly could, with the aid of meters and pen recorders, detect qualitatively unusual occurrences such as sudden shifts in a voltage or pressure, or oscillation about a given reading, or the gradual development of trends, the page-format CRT essentially has destroyed this capability.

As an example, observe the sequence of numbers in Figure 2. (This sequence might represent successive values of the last digit in a three-digit temperature reading.) It is virtually impossible in viewing such a sequence of changing numbers on a CRT, to detect the existence of a pattern. On the other hand, the same values plotted against time, as in Figure 3, begin to show a cyclic pattern.

Further data conditioning (such as performing autocorrelation and crosscorrelation) can result in extremely sensitive detection of cyclic disturbances. But even a plot of the moving average, as shown in Figure 4, clearly shows up the cyclic pattern. It is clear that detection of cyclic disturbances and second-order trends in sampled data can be enhanced by some display techniques and suppressed by others.

It appeared both feasible and desirable to develop display techniques for sampled data which would make use of man's superior pattern recognition capability. A number of promising pattern recognition display techniques to assist the test engineer in long-term monitoring of sampled data were considered. Autocorrelation, crosscorrelation, and cu-sum display techniques have already been described [1,4] as possible methods for detecting trends and departures from "stationarity" in a statistical sense.* A most

*This, of course, is the central problem in trend monitoring.

promising method is illustrated in Figure 5. Here we have plotted a "moving average" against a "moving standard deviation" (sample size = 5) for the sequence of numbers given in Figure 2. We have connected the dots with lines to show the time history of the sequence. The result is a spatially compressed, truncated time history of a sampled-data sequence which contains, in an extremely small display area, all of the pertinent statistical data inherent in the sequence. The computation of multiple moving averages and standard deviations, with a variable sample size, can be done by a display computer and displayed by CRT. Several dozen such plots on a single CRT display can be monitored simultaneously and easily by an observer feels that he could benefit from the additional information.

An alternative, simpler and in many ways superior method for computing moving estimates of the mean is the Exponentially Mapped Past.[5] The drawback of the method just described is that, in order to carry an n-sample moving estimate, n values must be preserved at all times in a "pipeline" storage (i.e., first-in, first-out). If many such moving estimates, corresponding to different parameters being monitored, must be carried, computer storage requirements may become excessive. To reduce the amount of storage required per moving estimate, it is better to use exponentially weighted estimates. In an exponentially weighted average, for example, a fixed fraction of the last value of the average is removed before adding the current sample. This has the effect of gradually decreasing the weight of the samples in the average as they age; in the ordinary moving average the aging occurs suddenly: the weight of a sample point goes from 1/n to zero when the sample becomes n+1 sampling intervals old. Rather than having to preserve n sample values, only one value, the exponentially weighted average, has to be preserved. This is called the Exponentially Mapped Past, or EMP, estimate of the mean.[5]

In order to fully specify any estimate of this type, a parameter denoting the influence of age on the weight of each sample point must be provided. In the case of the moving average, this parameter is the number of samples before the weight of a sample point is reduced by a factor of e. The reduction factor to be applied per sampling interval must be such that n consecutive reductions amount to a total factor of e; that is,

$$k = e^{-1/n} \tag{1}$$

for $n \gg 1$ we can write

$$k \simeq \frac{n-1}{n} \tag{2}$$

where k is the reduction factor per sampling interval, and n is the number of sampling intervals per time constant (n need not be an integer in the case of EMP averages).

As computed in formula (3) below, EMP averages with different time constants are properly scaled and therefore directly comparable in magnitude.

The EMP average is computed by means of the following expression:

$$\bar{X}_t = k\bar{X}_{t-1} + (1-k)x_t \qquad (3)$$

where $\bar{X}_t$ is the current value of the EMP average (available at time t); $\bar{X}_{t-1}$ is the previous value of the EMP average (available at time t-1); and x_t is the current value of the variable being considered.

Figure 6 shows a time history of raw data compared with the EMP estimate of the mean for n=6, plotted on the same time base. As in Figure 4, smoothing effect is again apparent.

In both the arithmetic and EMP techniques for obtaining running estimates of the mean, a capability for on-line variation in the weighting function is important. In a practical monitoring situation it is desirable to give the observer the capability to adjust sample size (or time constant) so that he can observe different phenomena. This is easily implemented (provided the input capability exists) by simply changing the appropriate parameter, k or n, in the computer.

EMP estimates of the standard deviation (σ) or the variance (σ^2) can be obtained as with the mean, by the following expression:

$$\sigma_t^2 = k\sigma_{t-1}^2 + (1-k)(\bar{X}_{t-x_t})^2 \qquad (4)$$

where σ_t^2 is the current value of the variance, σ_{t-1}^2 is the previous value, and k, $\bar{X}_t$ and x_t are as previously defined for the EMP estimate of the mean.

A plot of the EMP estimate of the moving variance (n=6), compared with the raw data from which it was computed, is shown in Figure 7.

Statistical Compression Display Format Evaluation

A number of techniques were evaluated for transforming and displaying a data stream in "compressed" format suitable for CRT display terminals. Earlier results had indicated that the time plot of a moving estimate of the mean was the most effective in making "visible" to an observer various kinds of disturbances in the data stream. The present study evaluated the time plots of other transformations as well as the other formats described in Reference [7].

The reader is now invited to review Figures 6, 7, and 8. Each is a computer print/plot of a sample of 100 points from a quasi-normally-distributed generating process with a mean of 28.0 (volts), and a standard deviation of 2.5 (volts). All "smoothing" (e.g., EMP estimates) uses n=6. At the 50th point the mean value in each figure changes to 30.0 (volts). The reader is invited to judge which transformation and display format makes this disturbance most visible and, hence, most useful to the test engineer.

Several general points are worthy of note. First, these figures are characteristic; other random number sets for the same conditions resulted in graphs quite similar to these. Secondly, the "hidden" disturbance used, was a change in the mean of 0.8σ as compared with 2.0σ used in the earlier study [6]; through proper choice of scale, larger changes would be more visible, and smaller changes less visible, regardless of the actual value of σ (as long as σ is at least 10 times the quantization error in the sampling system). Thirdly, an increase in N renders the change in mean more visible by smoothing, but increases the delay in seeing it.

Reviewing first Figure 6, it is almost impossible to detect the change at all, much less when it occurred, from the plot of raw data, x_t. (Experienced personnel could determine that a change had occurred by about the 75th point.)

From the EMP estimate of the mean in Figure 6, it is evident by about the 68th point that the mean has taken an excursion.

The EMP estimate of the variance, plotted in Figure 7, shows a confirmed change in the generating process by the 62nd point; as such, it is an excellent "flag," although it does not provide direct information on the nature of change.*

Because of the difficulty in portraying the characteristics of a dynamic real-time display in a static, single page, Figure 8 has been "doctored" to indicate the time sequence of the point positions. The sequence of points 21 - 40 should be disregarded since it represents the early growth of the EMP estimates of the mean and variance (see Figures 6 and 7). The sequence of points 61 - 70 reflects the delayed excursion in variance accompanying

*It does, however, provide indirect information from which it is possible to infer an excursion in the mean. The variance is defined as the first moment about the mean and is closely related to the concept of a "derivative." Since the first derivative of a step function is an impulse function, the presence of an "impulse" peak at the 64th point in Figure 7, with subsequent decay to normal, indicates strongly that a step change in the mean had occurred some time earlier. Similarly, a linear trend away from the mean would be indicated as a step change in the EMP estimate of the variance; this would help in identifying the early development of the trend.

the change in the mean. Note that all of these points are above the original mean of 28.0 (volts). The first excursion outside the "group" occurs at point 62, the next at 64; the sequence of points then remains outside the "group" with point 67.

Table I presents a comparative evaluation of the various formats in terms of detection delay. Similar results were obtained for other random number sets.

Although the scatter plot was found to be useless for aiding in detection of a shift in the mean, it has potential utility for obtaining simple and quick estimates of the autocorrelation of a sampled data stream.

Discussion

Most of the significant issues regarding these display techniques have already been covered in the preceding section. However, in addition to the data processing implications, which will be discussed in the next section, there is one additional aspect of the display problem which requires further discussion.

The problem is typified by the input requirements for existing detection techniques. To use them, the observer must select values for α (false alarm probability), for δ (size of change he is looking for), and n (sample size). Each has a critical effect on the detection delay; thus, although the techniques can be extremely sensitive when properly employed, the observer is required to (1) know precisely how (and why) to vary these parameters, and (2) know what in general to expect from the data stream in the way of disturbances. In general, this will be true of any such "rational" method which employs sophisticated statistical concepts. In addition, if nature of the disturbances is predictable, the entire technique can be programmed for automatic application by computer.*

In the real world of test and checkout, however, this is not the case. As pointed out in the Introduction, it is often difficult to specify precisely the kind or magnitude of disturbance for which a data stream is being monitored. Not only may the observer's expectation change from time to time, but also there will often be anticipated shifts (such as power supply loadings) which he should disregard. A "rational" method, then, would require either much more attention and adjustment on-line, or increased preplanning for automation.

*This was actually done for the Cu-Sum[8,9] technique to determine whether it was feasible for trained observers to "play with the parameters."

On the other hand, it was hoped that it would be possible to develop a data transformation and formatting method which would help an observer to "eyeball" the data, so to speak, and which would require a minimum of attention on the part of either himself or a computer. Apparently the Statistical Phase Plane Format, using EMP estimates of mean and variance, fulfills these requirements. For a given parameter, only two variables other than scale factors (the value of n, the exponential decay constant, and the number of points to be displayed) must be set, and then only infrequently. The observer can "eyeball" changes in mean and variance representing all of the categories of disturbance described earlier, and he can do it more easily and meaningfully using this format than using the separate time plots. Most important is the feature of a "visual centroid" in the Statistical Phase Plane, as indicated by the circle in Figure 8; normally, for a stable data stream all of the points in the display buffer (say, 25) can be made to appear on a small area of the CRT by appropriate selection of x- and y-scale factors. An excursion of the EMP estimate of the mean will quickly become evident as an excursion of a series of points from the heretofore stable "group," and will remain there for a fair amount of time.* In addition, after the individual excursions occur the "group" will move to a new position as new data points displace old ones.

The shape of the "group" also can have meaning. For example, a linear trend in the mean will cause a circular "group" of data points to increase in diameter and become slightly elongated along the variance axis. For a given scale, the size of the "group" can also be varied by changing the EMP decay constant, n, which will also affect detection delay. Its greatest effect, however, will be the visual stability of the "group." As an additional "eyeball" aid, a movable cursor symbol could be positioned to the centerof the "group" to provide a reference for changes in the mean and variance. A variable-diameter circle would be best for this purpose, since it could be made to circumscribe the "group" as in Figure 8.

Finally, several such groups, representing as many monitored parameters, can be monitored simultaneously on a single CRT.

It is most unfortunate that facilities could not be made available to this project for real-time simulation, using a CRT, of the techniques described. In spite of this major obstacle, we believe that the feasibility of the "Statistical Phase Plane" display has been demonstrated. Its operational advantages and disadvantages have yet to be adequately explored.

Conclusions

The work reported herein has explored the possibility of several methods for transforming and formatting measurement data,

*Thirty seconds, if the data rate is one per second and 30 points are maintained on the display.

so that an observer performing a monitoring function will be able to obtain the information he needs most readily in a manner which required the least amount of physical attention to the display. Our intent was to provide formats for CRT displays that will present more aspects of monitored data than is possible with the current tabular formats.

The results are evident to the reader in studying the figures. It would appear to be a fortunate combination of cir cumstances that the factors of format interpretability, flexibility, size, and ease of use,* as well as programming requirements and hardware constraints, all favor the Statistical Phase Plane Format using EMP estimates of the mean and variance. As noted in an earlier report [10] the Phase-Plane Format provides the greatest data compression in terms of display space, while also providing direct encoding of derivative information as a displacement on the display. It was this concept which suggested a parallel application to long-term monitoring. In the present instance, it is the statistical moment rather than the time derivative which appears to contain important anticipatory information concerning the stability of a data stream; hence the term "Statistical Phase-Plane."

*See reference [1] for detailed discussion of these factors.

References

[1] Wohl, J.G., Analysis of Man-Computer Roles in Prelaunch Fault Anticipation, Interim Report, 510-TM-12, SSD-65-192, Dunlap and Associates, Inc., Darien, Conn. June 1965.

[2] Wohl, J.G., Prelaunch Checkout Computers as Decision Aids-II. Data Conditioning and Display for Test Monitoring. 510-TM-15, SSD-65-226, Dunlap and Associates, Inc., Darien, Conn. Sept. 1965.

[3] Wiener, N., Extrapolation, Interpolation, and Smoothing of Sationary Time Series. MIT Press, Cambridge, Mass., 1948.

[4] Bendat, J.S. and A.G. Piersol, Measurement and Analysis of Random Data. John Wiley & Sons, Inc., New York, 1966.

[5] Otterman, Joseph. "The Properties and Methods for Computation of Exponentially Mapped Past Statistical Variables," IRE Trans. on Automatic Control, Vol. AC-5, No. 1, Jan. 1960, pp. 11-17.

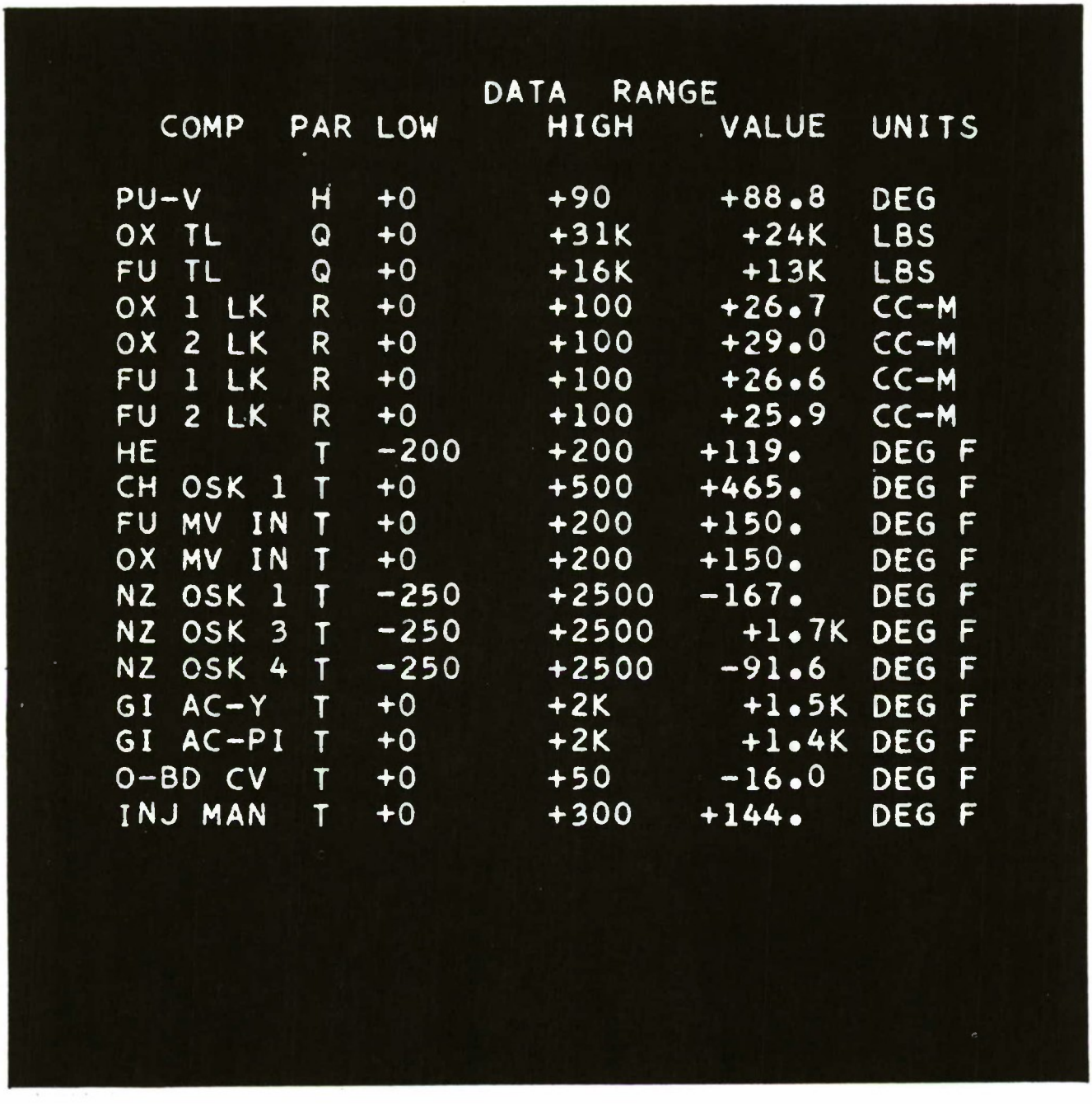
[6] Pepler, R.D., Data Conditioning and Display for Apollo Prelaunch Checkout: Multi-Parameter Monitoring Display, 675-TM-12, SSD-66-374, Dunlap and Associates, Inc., Darien, Conn. Dec. 1966.

[7] Wohl, J.G., Data Conditioning and Display for Apollo Prelaunch Checkout: Single Parameter Data Compression and Display, 675-TM-11, SSD-373, Dunlap and Associates, Inc., Darien, Conn. Dec. 1966.

[8] Johnson, N.L., "A Simple Theoretical Approach to Cumulative Sum Control Charts," J. of American Statistical Assn., 56, 1961.

[9] Duncan, A.J., Quality Control and Idustrial Statistics, R. D. Irwin, Inc., 1965.

[10] Affinito, F.J. and J.G. Wohl, Data Conditioning and Display for Apollo Prelaunch Checkout-Servo Response Display Technique, 675-TM-7, SSD-66-353, Dunlap and Associates, Inc., Darien, Conn. Dec. 1966.



COMP	PAR	DATA RANGE LOW	HIGH	VALUE	UNITS
PU-V	H	+0	+90	+88.8	DEG
OX TL	Q	+0	+31K	+24K	LBS
FU TL	Q	+0	+16K	+13K	LBS
OX 1 LK	R	+0	+100	+26.7	CC-M
OX 2 LK	R	+0	+100	+29.0	CC-M
FU 1 LK	R	+0	+100	+26.6	CC-M
FU 2 LK	R	+0	+100	+25.9	CC-M
HE	T	-200	+200	+119.	DEG F
CH OSK 1	T	+0	+500	+465.	DEG F
FU MV IN	T	+0	+200	+150.	DEG F
OX MV IN	T	+0	+200	+150.	DEG F
NZ OSK 1	T	-250	+2500	-167.	DEG F
NZ OSK 3	T	-250	+2500	+1.7K	DEG F
NZ OSK 4	T	-250	+2500	-91.6	DEG F
GI AC-Y	T	+0	+2K	+1.5K	DEG F
GI AC-PI	T	+0	+2K	+1.4K	DEG F
O-BD CV	T	+0	+50	-16.0	DEG F
INJ MAN	T	+0	+300	+144.	DEG F

FIGURE 1. **Typical** Tabular or "Page" CRT Display Format for **Centralized** Test Monitoring during Apollo/Saturn Prelaunch Checkout

Component tested, parameter type, low and high limits, instantaneous value, and units of measurement are shown. An out-of-limits indication (e.g., second line from bottom) is called to the test engineer's attention by periodic interruption of the appropriate line of information, causing that line to "flash."

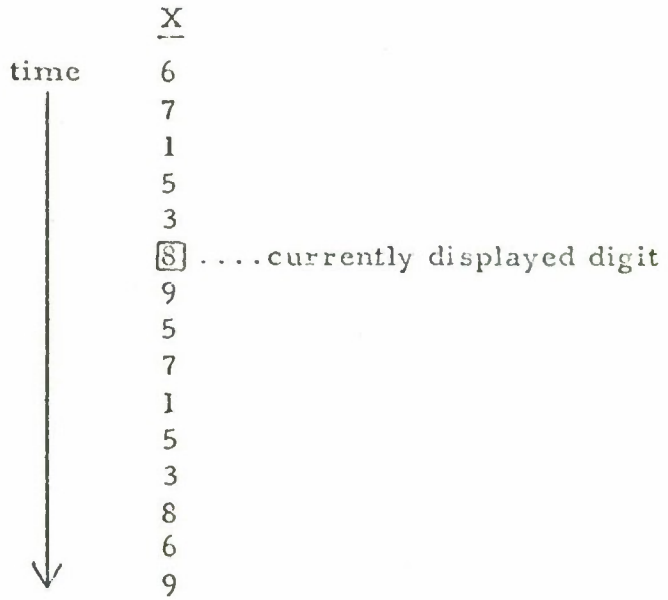


FIGURE 2. Typical Time Sequence of Digits Appearing in a Single Character Space on a Page-Format CRT Display Such as that Shown in Figure 1.

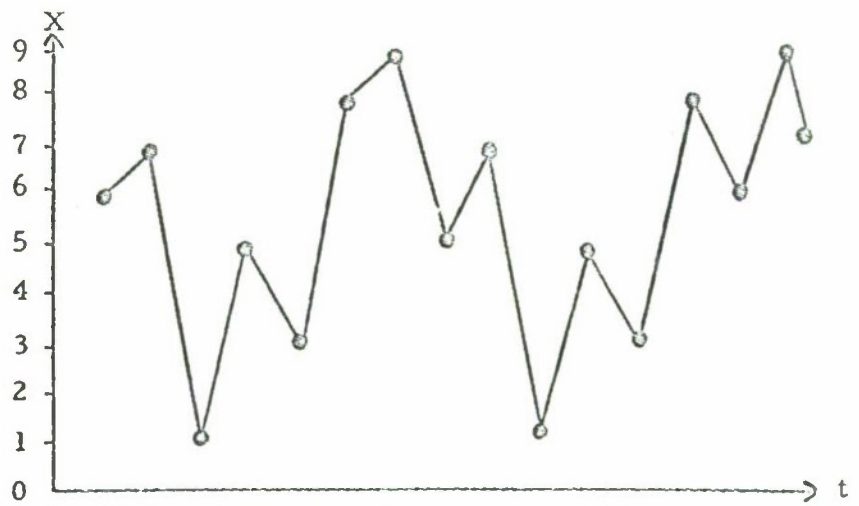


FIGURE 3. Same Sequence Plotted vs. Time in Graph Format.

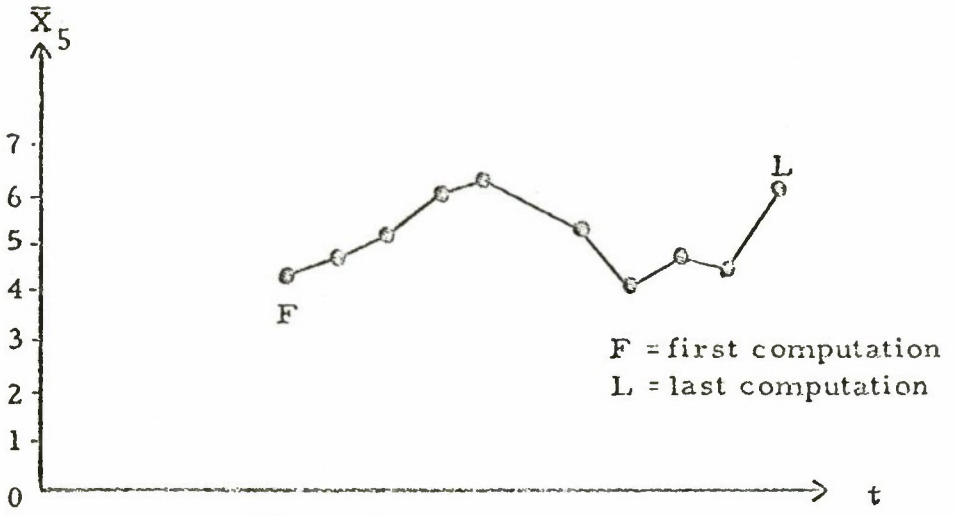


Figure 4.

Plot of "Moving Average" vs. Time of Sequence in Figure 2. Sample Size = 5.

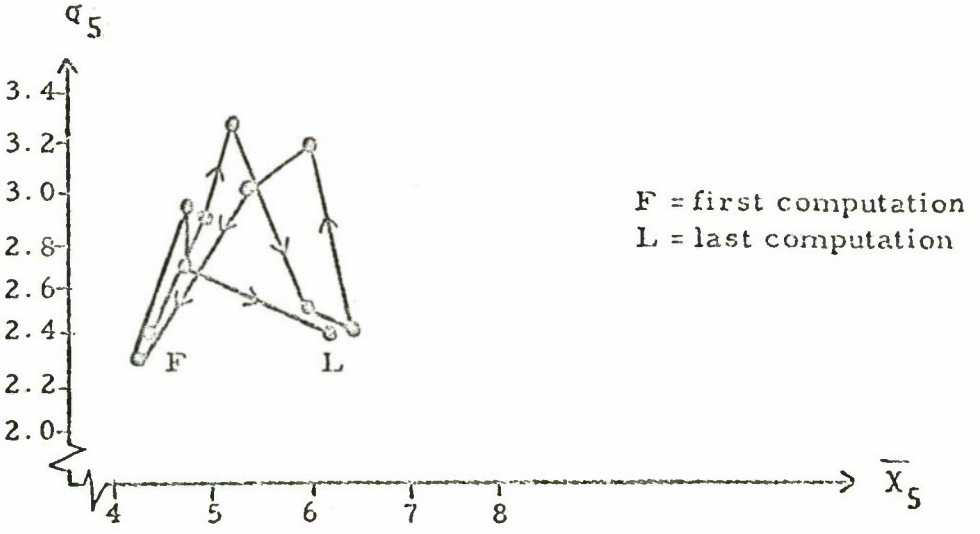


Figure 5.

Plot of "Moving Average" vs. "Moving Standard Deviation" for Sequence in Figure 2, Indicating Data Compression. Sample Size = 5.

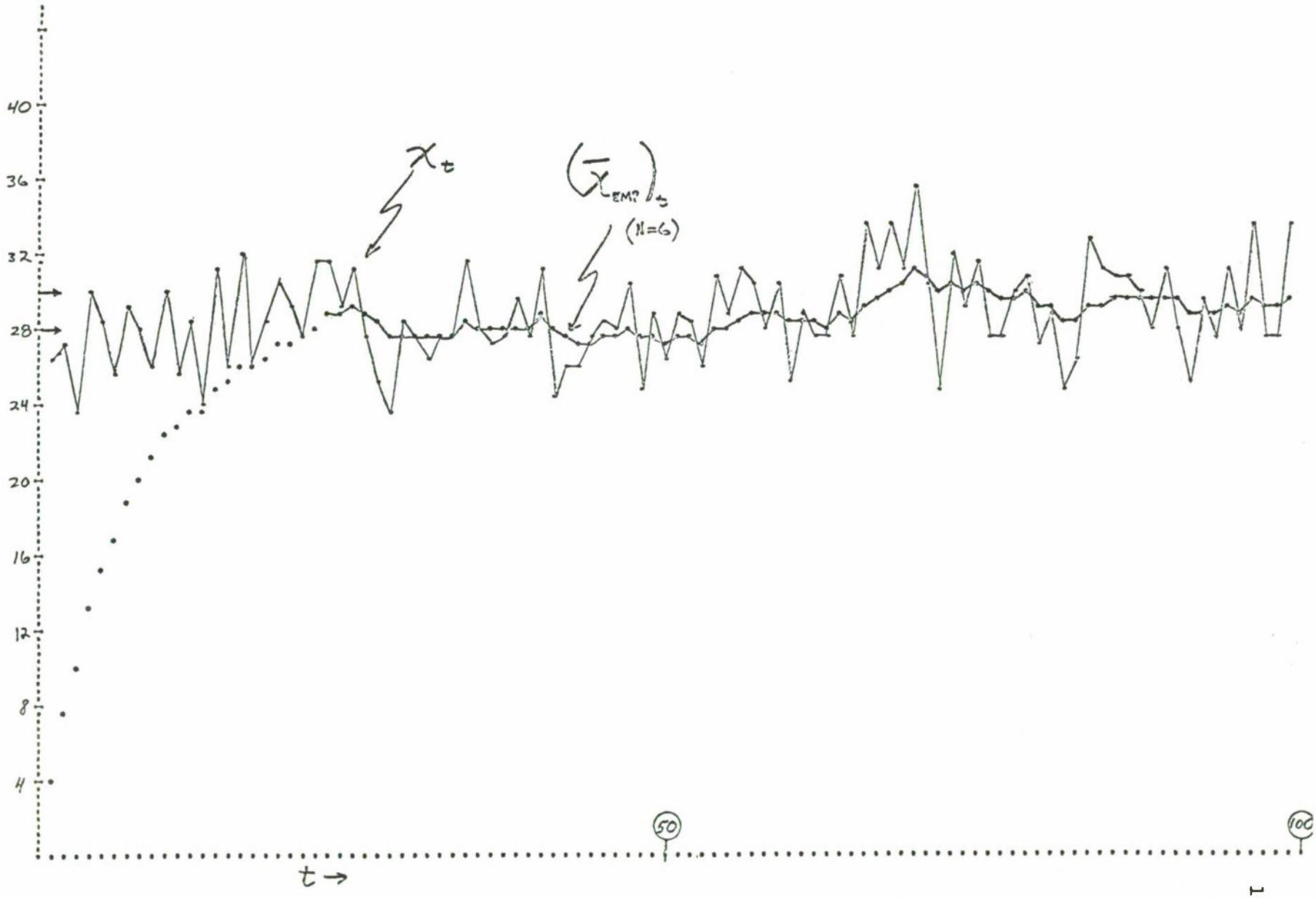


Figure 6. Time plot of EMP moving estimate of the mean for N=6, superimposed upon the raw data samples.

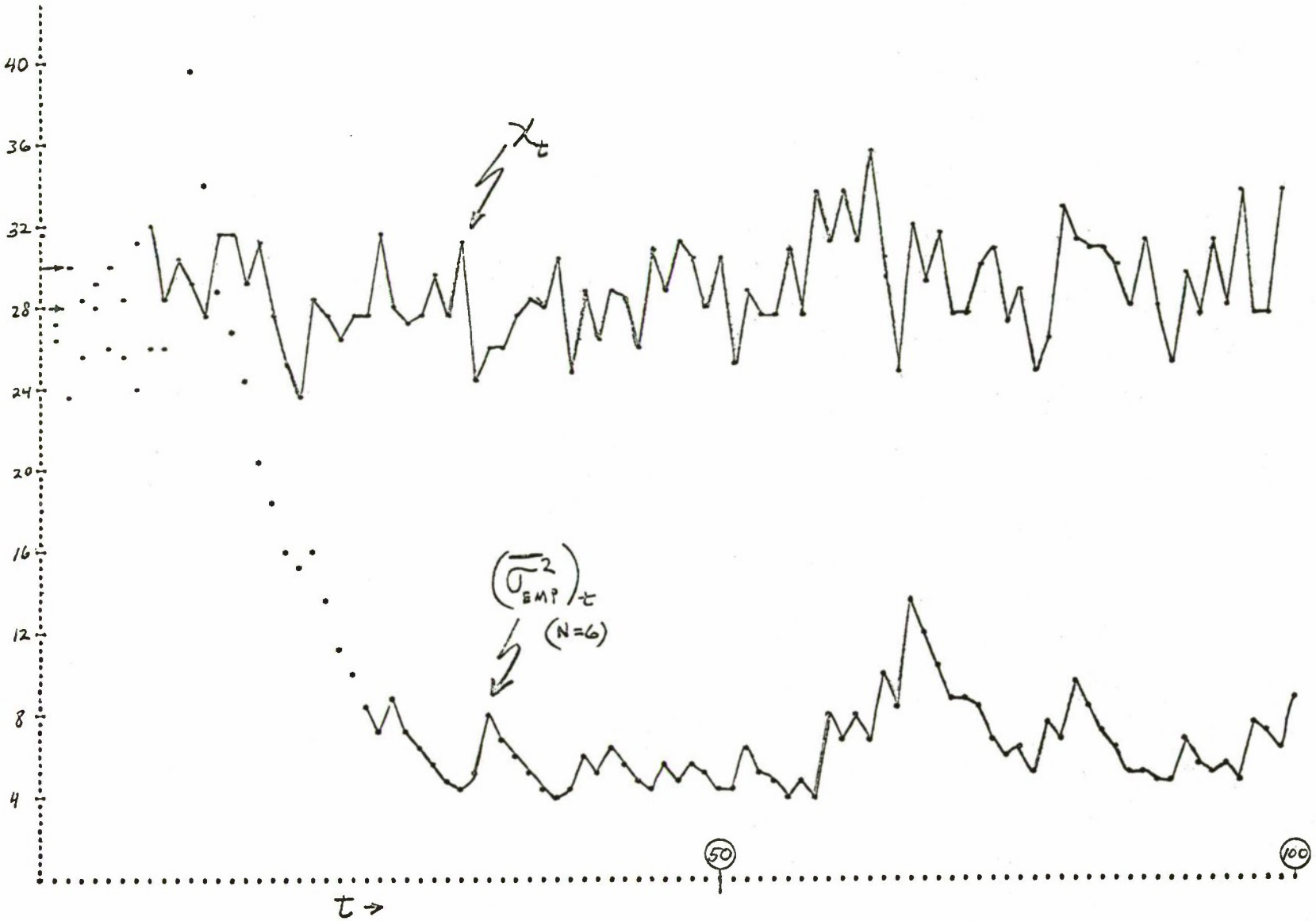


Figure 7. Time plot of EMP moving estimates of the variance, with raw data samples shown for reference.

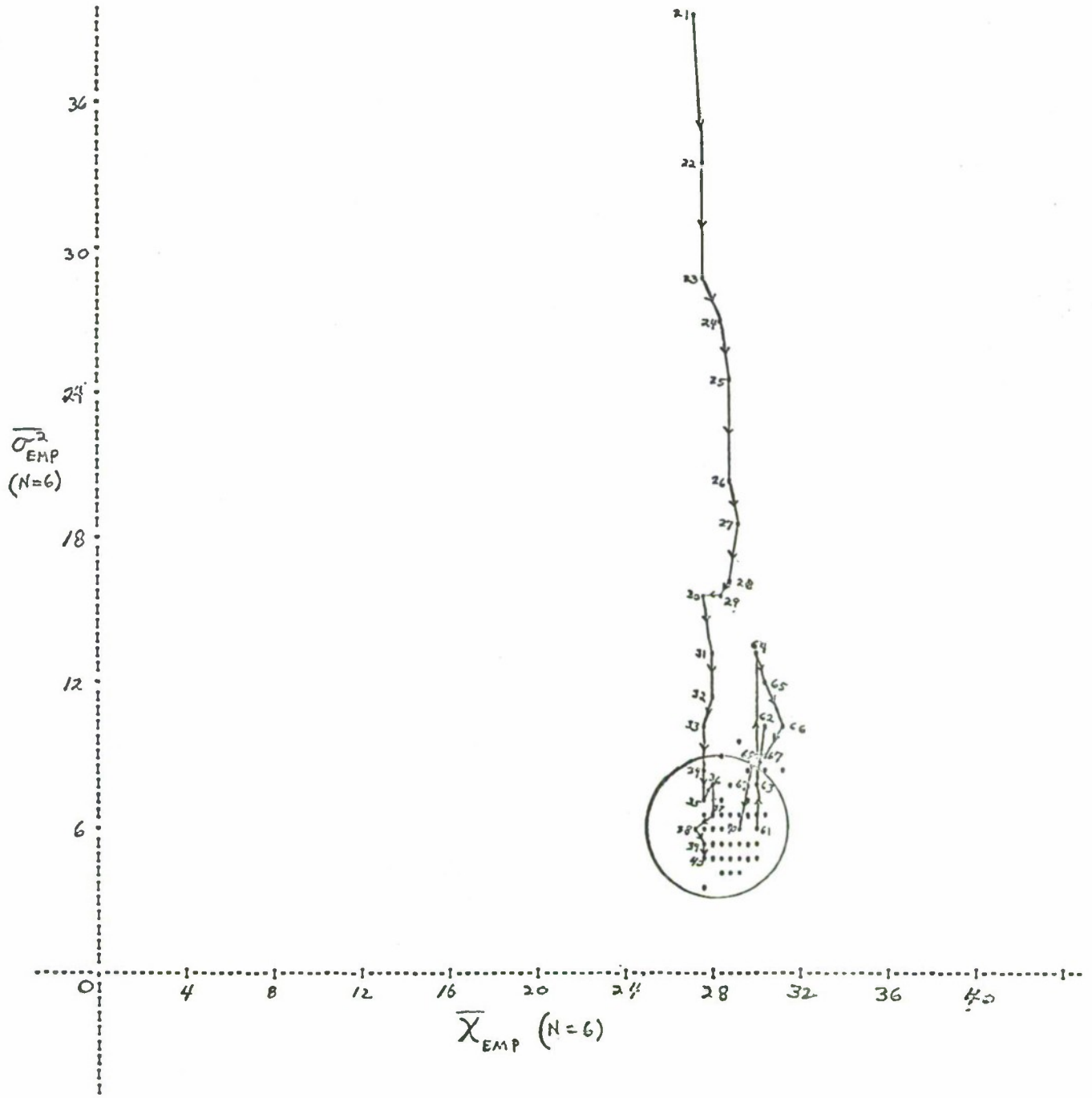


Figure 8. Statistical Phase-Plane Format, or plot of EMP moving estimate of the mean against the EMP moving estimate of the variance.

FORMAT AND TRANSFORM	DETECTION DELAY (No. of Samples or Units of Time)
Time Plot	
Raw data	25-point delay
First Derivative	(useless)
Arithmetic Mean	18-point delay
EMP Mean	18-point delay
EMP Variance	12-point delay
Statistical Phase Plane	
EMP Mean vs. EMP Variance	12-point delay*
Cu-Sum Plot*	
$\delta = 2.0$	36-point delay
$\delta = 0.8$	12-point delay*

TABLE I

*Note that use of the Cu-Sum technique gave results equivalent to that of the Statistical Phase Plane. However, it requires that three parameters instead of one be set beforehand, and that it be "tuned" to the process. A wrong selection (see Table for example) renders the method useless.

PROGRESS IN QUANTIFICATION OF MUSCULAR ACTION

Julia T. Apter, M.D., Ph.D.
Division of Surgery, Presbyterian-St. Luke's Hospital
Department of Surgery, University of Illinois
Chicago, Illinois
and
William W. Graessley, Ph.D.
Department of Chemical Engineering, Northwestern University
Evanston, Illinois

In March, 1969, at this conference, we presented a model for muscular behavior that promised to be useful in quantifying muscular action for the purpose of analyzing the capabilities of handicapped individuals and in developing prostheses (Apter and Graessley, 1970). This paper gives a progress report of further work on the model which qualitatively fit a wide variety of muscular behavior, but which had not been tested quantitatively.

Since last year, five big steps have been taken in making the model useful; (1) The manuscript was accepted for publication in the Biophysical Journal; (2) Development of a physical basis for the dependence of model parameters E_1, E_2 and η on the configuration of macromolecules of muscle to replace the empirical relation previously used; (3) Further experimentation and computer analysis of the model to show its range of quantitative validity (Figures 1, 2, 3, 4); (4) Securing an instrument capable of measuring stress, strain and strain rate in living human subjects and power to quantify human behavior (Taylor and Brown, 1965).

This progress has been facilitated by H. D. Landahl, J. Provan, L. Amiot, R. Taylor, T. Brown and the Cornell Aeronautical Laboratories.

Supported in part by USPHS Grant number GH-14659-03 from the National Institute of General Medical Sciences, by the Physics Division of the Atomic Energy Commission and by the Adam Birli Computing Fund.

The events known as the "sliding filament hypothesis" of muscular contraction (Huxley, 1957, 1963) were incorporated empirically into the model. The rearrangement of actin and myosin that takes place as muscle shortens (or as the unstrained length ℓ_o decreases) were assumed also to be associated with increases in model parameters E_1, E_2 and η. Their interrelation was compiled from experiments on smooth muscle (Apter et al., 1966, 1968a, 1968b).

The hyperbolic form of the relationship was arbitrary, and chosen for simplicity. An exponential or even a linear (rather than hyperbolic) relation would also have given a reasonable representation of the data. Some justification for the particular form selected, however, can be elicited from the following argument (Apter, 1961).*

Suppose the stimulant, N, is taken to be a single factor which acts on a single muscle component A. Consider an equilibrium between N, free activator A, and a bound activator NA:

$$N + A \rightleftharpoons NA \tag{1}$$

with an equilibrium constant, K. The equilibrium activities (A) and (NA) are

$$(A) = A_o/(1 + kn) \tag{2}$$

$$(NA) = kn\ (A) = kn\ A_o/(1 + kn) \tag{3}$$

with $k \equiv K^{-1}$ and n = the concentration of N. A_o is the total available amount of activator. Let the elastic parameters E (E_1 or E_2 or ℓ_o or η) be a linear combination of (A), the free and (NA) the bound activator; i.e.,

$$E = \alpha_4\ (A) + \beta_4\ (NA) = \frac{\alpha_4\ A_o}{1 + kn} + \frac{\beta_4\ kn\ A_o}{1 + kn} \tag{4}$$

or

$$= \beta_4 A_o + \frac{(\alpha_4 - \beta_4)A_o}{1 + kn} \tag{5}$$

Comparing (5) with equations for E in previous paper we see that

$$E^o = \alpha_4 A_o \quad ; \quad E^\infty = \beta_4 A_o \tag{6}$$

where E^o is in the relaxed state and E^∞ is in the contracted.

*Suggested by Herbert Landahl, Professor, University of California, San Francisco.

The simplest model, then, would have

$$k_1 = k_4 = k_5 = k_6 = K^{-1} \tag{7}$$

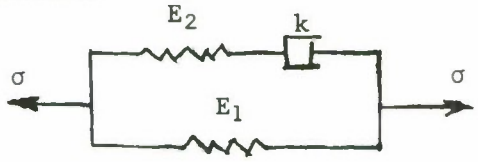
and therefore all the k_i would be related to the equilibrium constant of equation (1).

Concerning the equation of motion of these parameter models

$$\sigma + \frac{\eta}{E_2}\dot{\sigma} = E_1\varepsilon + (E_1 + E_2)\frac{\eta}{E_2}\dot{\varepsilon} \tag{8}$$

This equation is strictly valid for time independent material parameters. However, from the equations (1) to (7) and (17) below, these material parameters depend on the concentration of substance N, namely n, which in turn satisfies your time differential equation (3), and hence cannot be assumed time independent.

We propose the following equation (16), whose derivation we now present, as an alternative.*



The total stress σ is split between the Maxwell unit and the spring as follows:

$$\sigma = \sigma_M + \sigma_S \tag{9}$$

where σ_M is the stress in the Maxwell unit's arm and σ_S the stress in the spring's arm. The stress in the spring is related to the total strain of the model via:

$$\sigma_S = E_1\varepsilon \tag{10}$$

However, this same strain comprises of the two strains across the Maxwell unit so that:

$$\varepsilon = \varepsilon_{E_2} + \varepsilon_\eta \tag{11}$$

where ε_{E_2} is the strain in the Maxwell unit's spring and E_η the Maxwell unit's dashpot. They are related to σ_M through the equations:

*Suggested by James Provan Research Associate and D. R. Axelrad, Professor, Micromechanics Laboratory, McGill University.

$$\varepsilon_{E_2} = \frac{1}{E_2}\sigma_M \quad , \quad \dot{\varepsilon}_2 = \frac{1}{\eta}\sigma_M \tag{12}$$

Now, we just have to eliminate σ_M, σ_S, ε_{E_2} and ε_η from (9) - (12) in the following manner in order to arrive at a governing differential equation.

Substitute (12) into (11):

$$\dot{\varepsilon} = \dot{\varepsilon}_{E_2} + \dot{\varepsilon}_\eta = \frac{d}{dt}\left(\frac{1}{E_2}\sigma_M\right) + \frac{1}{\eta}\sigma_M \tag{13}$$

Combine σ_S of (10) in a manner similar to the right hand side of (13):

$$\frac{d}{dt}\left(\frac{1}{E_2}\sigma_S\right) + \frac{1}{\eta}\sigma_S = \frac{d}{dt}\left(\frac{E_1}{E_2}\varepsilon\right) + \frac{E_1}{\eta}\varepsilon \ , \tag{14}$$

and add (13) to (14) to obtain the differential equation:

$$\frac{d}{dt}\left(\frac{E_1}{E_2}\varepsilon\right) \frac{E_1}{\eta}\varepsilon + \dot{\varepsilon} = \frac{d}{dt}\left(\frac{1}{E_2}\sigma\right) + \frac{1}{\eta}\sigma \quad , \tag{15}$$

where use has been made of (9). If we assume the material parameters to be time independent then (15) reduces to (8), or our previous equation (1). If, however, they are not time independent then (15) may be written as:

$$\frac{d}{dt}\left(\frac{\sigma}{E_2}\right) + \frac{\sigma}{\eta} = \left\{\frac{E_1}{E_2} + 1\right\}\dot{\varepsilon} + \left\{\frac{d}{dt}\left(\frac{E_1}{E_2}\right) + \frac{E_1}{\eta}\ \varepsilon\right\}. \tag{16}$$

The solution of this equation is not appreciably different from our previous analysis, but rests on firmer ground and makes quantification more rigorous.

In our previous paper we point out that the values of the constants E_1, E_2 and η depend on the macromolecular arrangement which in turn is influenced by some substance N near the intracellular macromolecules. It has been assumed that the concentration of N varies with time according to

$$\dot{n} = k_2\varepsilon - k_3 n + S(t) \tag{17}$$

$$= k_2\left[\frac{\ell(1 + k_1 n)}{\ell_o^o + \ell_o^\infty k_1 n} - 1\right] - k_3 n + S(t) \quad , \tag{18}$$

i.e., $\dot{n} = \dot{n}[n,\ell]$. Furthermore, the rest length and the viscoelastic material parameters of the muscle have been assumed to be simple functions of the concentration n as already summarized in (5).

Hence, we see that equation (16) depends implicitly on the rate of concentration of N. We suggest two ways of handling this problem:

(i) assess the values of these material parameters as functions of the concentration using analogue computational methods before substituting into equation (16). Hence, (16) will be an equation between ℓ, the current length and σ the stress on the muscle as a function of the concentration n and the time t.

(ii) substitute the indicated values of the unstretched length and material parameters, as functions of n, into the dynamic equation of state (16) which would lead to a relation of the form:

$$\left\{\frac{(E_1^o + E_1^\infty k_4 n)(1 + k_5 n)}{(E_2^o + E_2^\infty k_5 n)(1 + k_4 n)} + 1\right\}\left\{\frac{1}{\ell_o^o + \ell_o^\infty k_1 n}\left\{\dot{\ell}(1 + k_1 n) + \ell\dot{n}\left[k_1 - \frac{\ell_o^\infty k_1 (1 + k_1 n)}{\ell_o^o + \ell_o^\infty k_1 n}\right]\right\}\right\}$$

$$+\left\{\dot{n}\left\{\frac{(E_1^\infty k_4 + E_1^o k_5 + 2E_1^\infty k_4 k_5 n)}{(E_2^o + E_2^\infty k_5 n)(1 + k_4 n)} - \frac{(E_1^o + E_1^\infty k_4 n)(1 + k_5 n)(E_2^\infty k_5 + E_2^o k_4 + 2E_2^\infty k_4 k_5)}{[(E_2^o + E_2^\infty k_5 n)(1 + k_4 n)]^2}\right\}\right.$$

$$\left. + \frac{(E_1^o + E_1^\infty k_4 n)(1 + k_6 n)}{(n^o + n^\infty k_6 n)(1 + k_4 n)}\right\}\left[\frac{\ell(1 + k_1 n)}{\ell_o^o + \ell_o^\infty k_1 n} - 1\right] \qquad (19)$$

$$= \frac{1}{E_2^o + E_2^\infty k_5 n}\left\{\dot{\sigma}(1 + k_5 n) + \sigma\dot{n}\left[k_5 - \frac{E_2^\infty k_5 (1 + k_5 n)}{E_2^o + E_2^\infty k_5 n}\right]\right\} + \frac{\sigma(1 + k_6 n)}{n^o + n^\infty k_6 n} \quad .$$

This relation expresses a linear viscoelastic physiological model appropriate for muscle behavior and indicates clearly the dependence on the concentration n. Using (18) simultaneously with (19), then again ℓ and σ are related as functions of n and t.

Both of these approaches are handled readily by computation and lead to estimation of parameters with methods that have been useful in determining molecular orbitals.

The methods rely on a package of general-purpose Fortran subroutines for solving problems in non-linear least squares, non-linear regression, or non-linear estimation: that is, problems of fitting experimental data to a model, by seeking the "best" values of certain parameters which occur non-linearly in the model. In mathematical terms, the programs seek a local minimum (or maximum) of any unconstrained function of k variables which can be expressed in some natural way as a sum of r squares, where $r \geq k$. In particular, the programs can be used to find a local solution to a system of k simultaneous non-linear equations in k unknowns.

The general-purpose program for this purpose was written for a modern computer by Taylor and Brown (1961). It differed from previous such programs in four main respects: it was Fortran-oriented; it employed a mixture of the Newton-Raphson-Gauss algorithm used previously with a steepest-descent technique; it required the user to provide partial derivatives of the model as well as function values, thereby encouraging the use of analytic methods of differentiation in preference to finite differences; and it accepted responsibility for coping in some way with any problem that might be presented to it, however ill-conditioned.

The program has subsequently been applied to a wide variety of model-fitting problems in physics, chemistry, and engineering. Many refinements have been incorporated as a result of experience with these applications. In this biological problem, we have found that the "goodness of fit" is strongly dependent on having a good set of initial estimates of the parameters. Our problem now lies in finding the experiments most likely to provide good initial estimates on humans without causing fatigue.

The results depicted in Figures 1, 2, 3 and 4 were obtained using such parameter estimations techniques, performed by F. Morehouse and L. Amiot of Argonne National Laboratories.

REFERENCES

Apter, J. T. 1961, Amer. J. Ophthal. 51:269.

Apter, J. T. and W. W. Graessley, 1970. NASA-Univ. Conf. on Manual Control. 215:523-540.

Apter, J. T. and Marquez, E. 1968a Circ. Res. 22:393.

Apter, J. T. and Marquez, E. 1968b Biorheology 5:285.

Apter, J. T., Rabinowitz, M. and Cummings, D. H. 1966. Circ. Res. 19:104.

Fenn, W. O. and Marsh, B. S. 1935. J. Physiol. 85:277.

Huxley, H. E. 1957. Prog. Biophys. 7:255.

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.

Staverman, A. J. and Schwarzl, R. 1956. Linear Deformation and Behavior of High Polymers in Die Physik der Hochpolymeren IV Ed. H. A. Stuart, Springer-Verlag, Berlin.

Taylor, R. I. and Brown, T. H. 1965. J. Phys. Chem. 69:2316.

Table 1

Modelled Material	ℓ_o^o	ℓ_o^∞	k_1	k_2	k_3	k_4	k_5	k_6	E_1	E_1^o	E_1^∞	E_2	E_2^o	E_2^∞	η	η^o	η^∞
	(units of time are seconds)								(dynes cm ^{2} x 10^{-5})						(poise x 10^{-3})		
"Elastin"			0	0	0	0	0	0	50			0.5			10		
"Relaxes smooth muscle"	1.85		0.1	0.2	0.8	0.1	0.1	0.1		10			1			10	
"Contracted smooth muscle"	1.85		0.1	0.5	0.4	0.1	0.1	0.1			50			7			160
	1.70		0.07	0.4	0.2	0.07	0.07	0.07			5.9			7.6			181
"Striated muscle"	2.0		0.75	.9	.15	1.0	1.0	1.0		10	50		1	7		10	160

Superscript o refers to "relaxed muscle."

Superscript ∞ refers to "contracted muscle."

Figure 1. Experimental data (circles) obtained from canine urinary bladder 24 hours after removal from the dog and with demonstrably dead muscle, since it did not respond to a stimulant. Phase shift ϕ and absolute dynamic modulus [E] are plotted as functions of frequency of induced oscillation (strain). The match of data and model curves is well within experimental error (Apter et al., 1966). This is not surprising since the stress relaxation curve of this specimen was clearly exponential so that parameter estimation techniques gave reliable values for E_1, E_2 and η with all $k_i = 0$ (Apter and Marquez, 1968b). See equations 1 to 7. Therefore, the specimen acted like elastin, an inert polymer (Staverman and Schwarzl, 1956) and matched model behavior (triangles) very well.

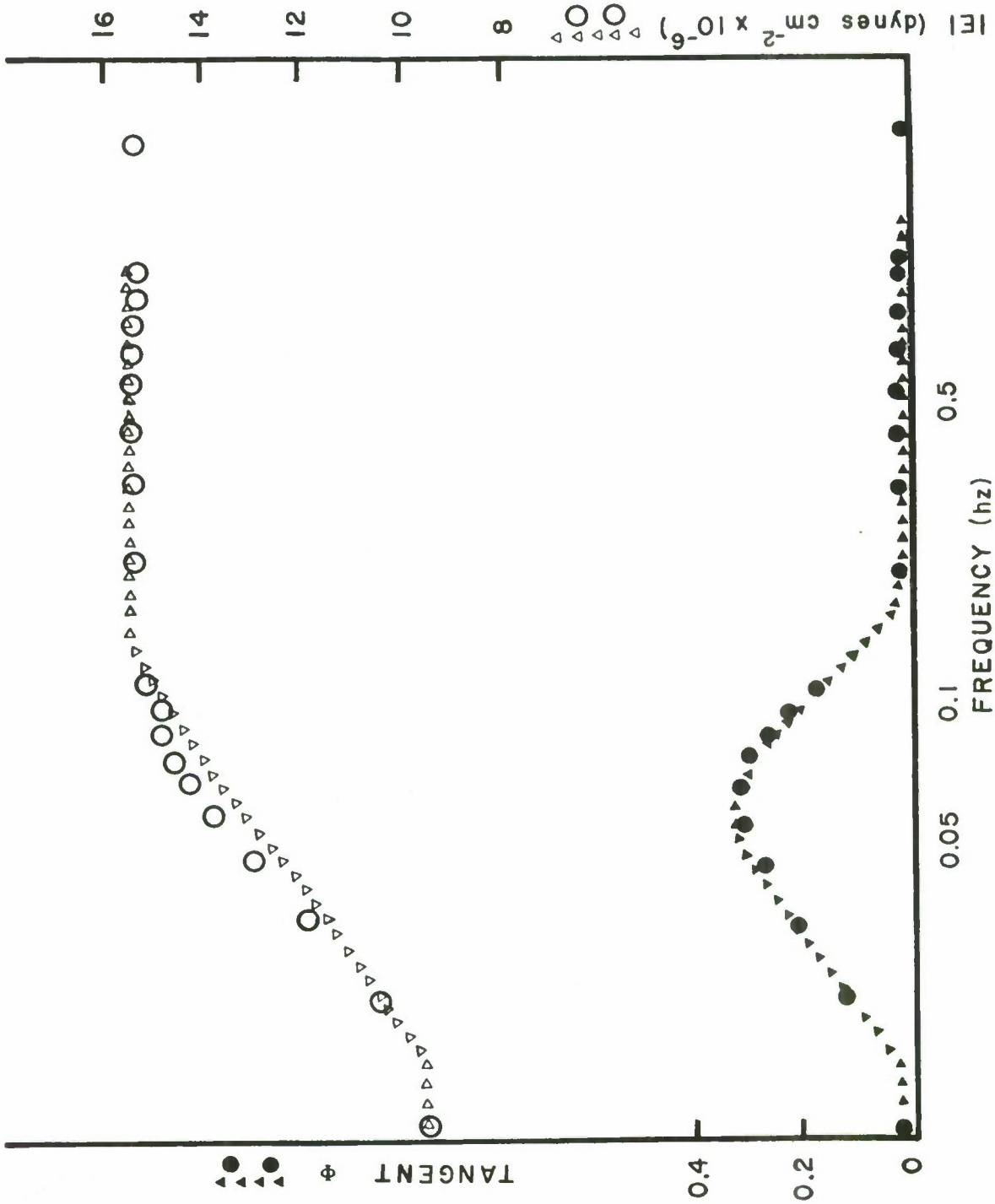


Figure 2. Experimental data (circles) obtained from canine urinary bladder immediately after removal from dog. Analysis of results is as in Figure 1. The phase lead at low frequencies changes to the customary phase lag at precisely the same frequency as the model (triangles) provided the E_1, E_2 and η as well as the k_i ($i = 1 \cdots 6$) were obtained by parameter estimation techniques from stress relaxation curves. Generally, the peak stress gave $E_1 + E_2$ and the amount of recontraction gave k_2/k_3. This fit is satisfactory and represents behavior unique for living polymers like muscle.

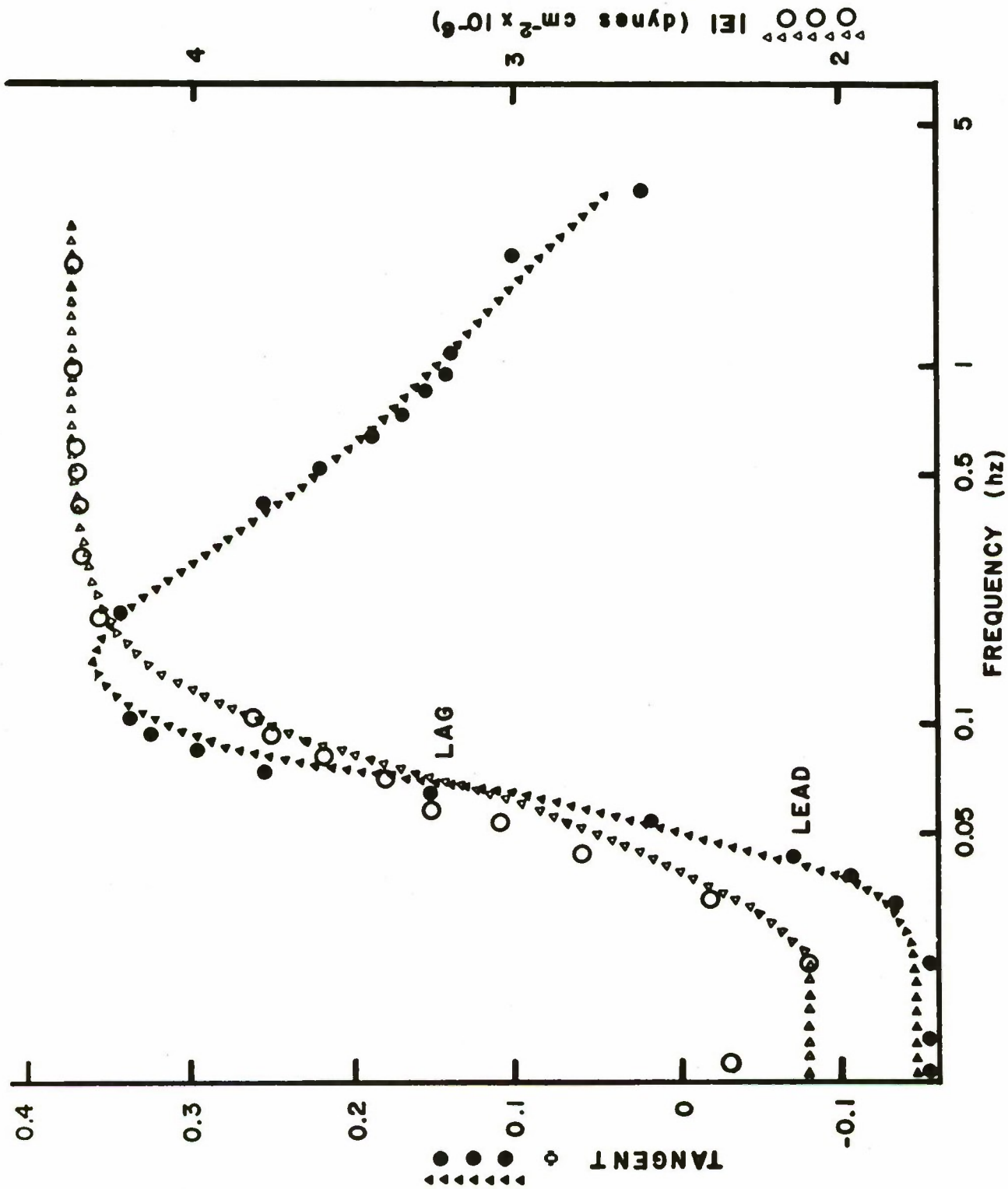


Figure 3. The experimental data (circles) were obtained from the same specimen as in Figure 2, but after the addition of sufficient urecholine to raise the mean tension above the in vivo level and to reduce the lumen circumference. The stress and strain were computed in accordance with the new dimensions. Analysis of results is as in Figure 1. The modulus is generally higher and the phase lead angle is lower at low frequencies than it is for more relaxed smooth muscle. The k_i ($i = 1 \cdots 6$) are listed in Table 1 and were computed from stress relaxation curves. These are responsible for this behavior which is unique for active muscle. The computer-generated solution of the model incorporating these parameter values are plotted as triangles.

A similar fit was obtainable for striated muscle studied immediately after removel from a dog at 0° C (see Figure 4). However, the onset of rigor mortis (characterized by a large phase lead and ever increasing modulus) made it impossible to be certain of steady state conditions. The deviations from a good fit in these data probably indicate inadequacies in the parameter estimation techniques used here.

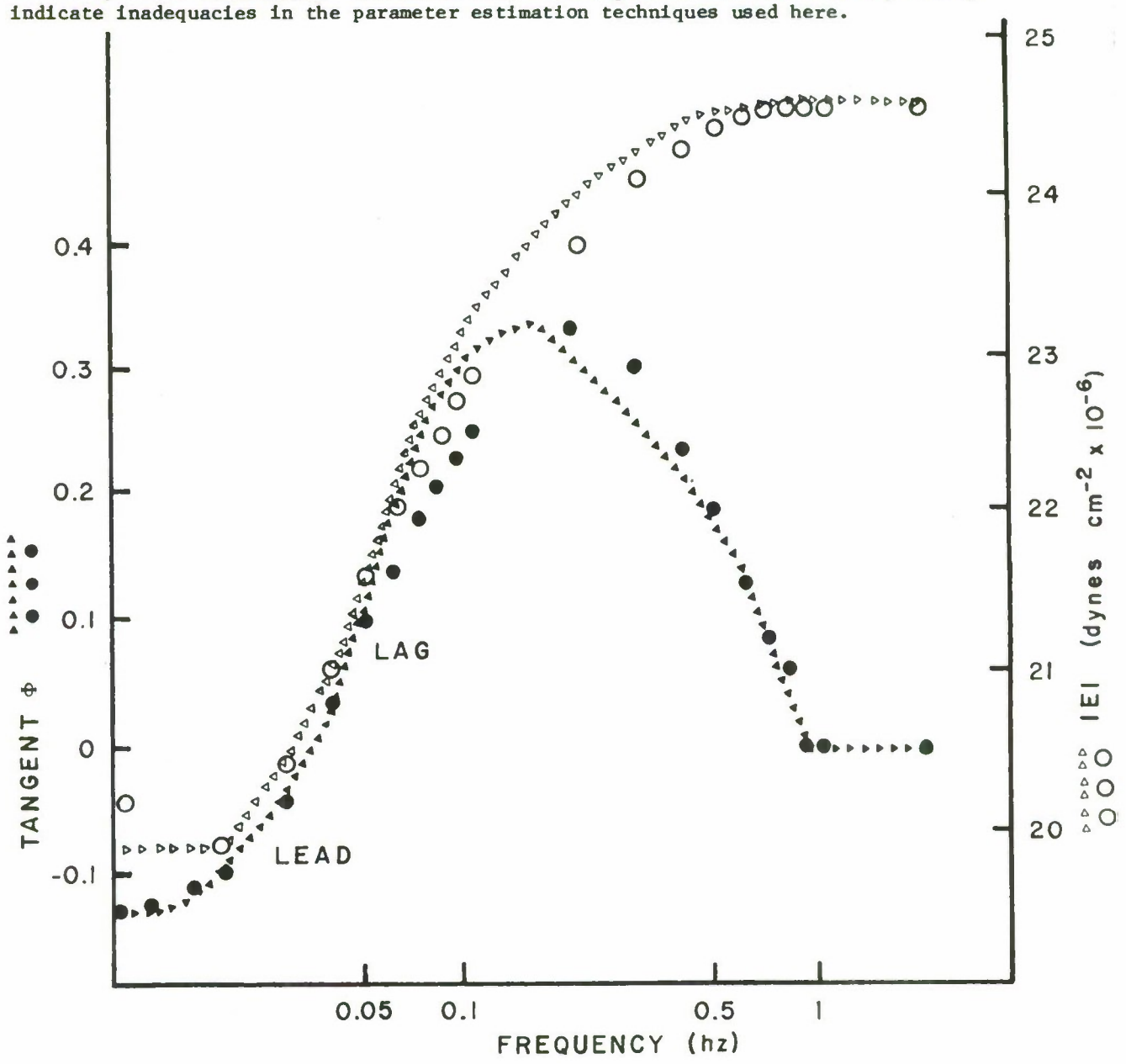
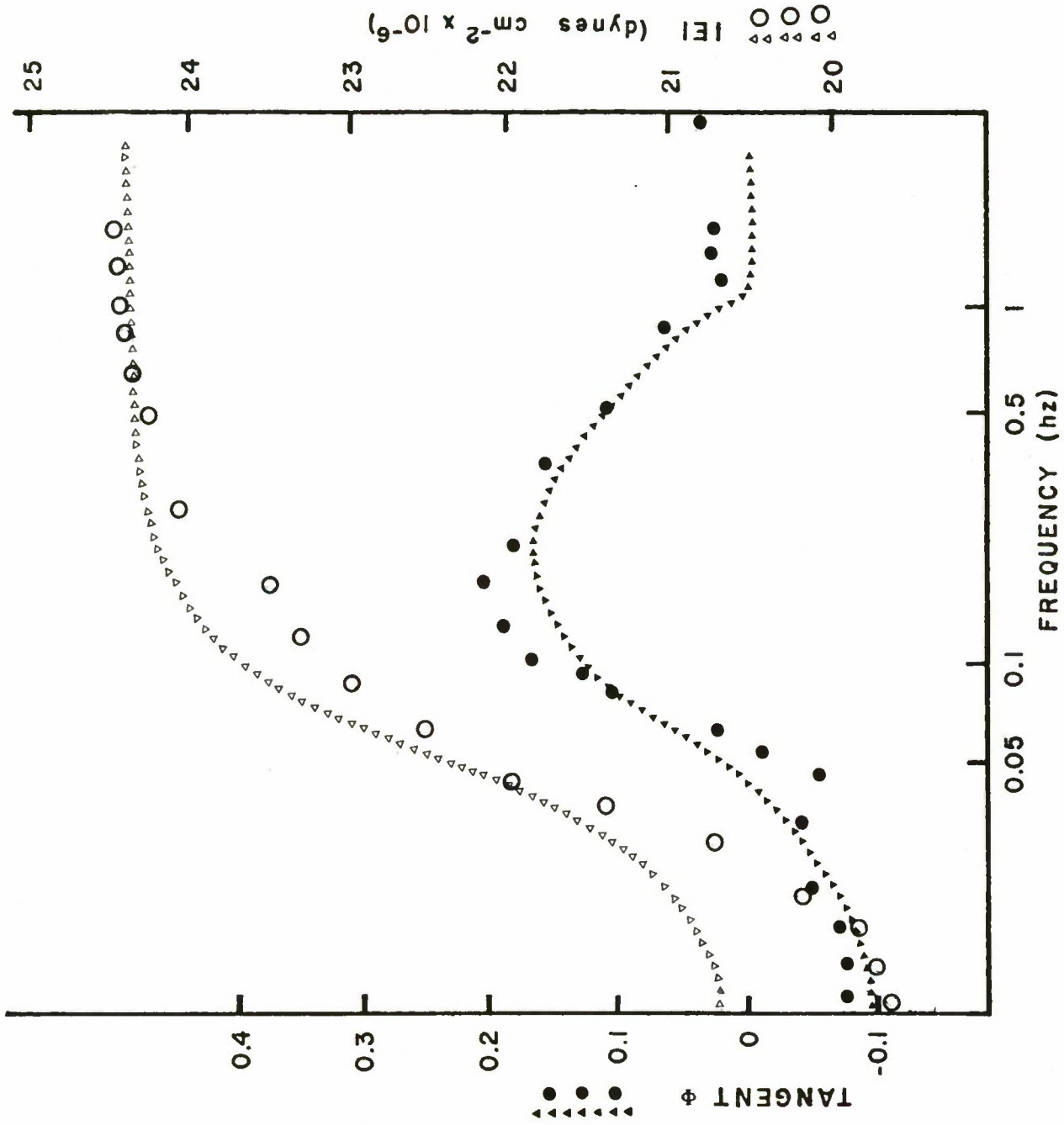


Figure 4. Data (circles) obtained from cat lateral rectus muscle, a striated muscle having all fibers of equal length and only four or five muscle fibers thick. Triangles are output from computer. This fit is not nearly so good as obtained from smooth muscle.



A MODEL OF MUSCLE ACTIVITY

Gerald L. Gottlieb and Gyan C. Agarwal
Biomedical Engineering Department
Presbyterian-St. Luke's Hospital
Chicago, Illinois 60612
and
Colleges of Medicine and Engineering
University of Illinois, Chicago, 60680

ABSTRACT

Human skeletal muscle may be regarded as an electro-mechanical transducer. Its physiological input is a neural signal originating in the spinal cord and its output is force and muscle contraction, these both being dependent on the mechanical load.

The paper describes some experimental data taken during various voluntary efforts of the foot. Recordings were made of the soleus and anterior tibial muscles' electromyograms and of the foot torque. By various filtering and averaging techniques, "average" responses during step and ramp efforts of various duration are derived.

Some of the above experiments are simulated by an analog model, the input of which is recorded physiological soleus muscle electrical activity. The output is simulated foot torque. The results are discussed and a model variable which may be identified with the muscle "active state" is described.

This work was partially supported by NIH Training Grant and PSL grant-in-aid.

Introduction

Associated with the voluntary contraction of normal skeletal muscle are two easily detected phenomena. Most important from a utilitarian point of view is the force that the muscle generates. Preceeding and concurrent with such a mechanical manifestation there is an electrical response from the muscle which may be measured by needle or surface electrodes and is called the electromyogram or EMG. This EMG is an interference pattern formed by the action potentials of many muscle fibers firing in response to their motoneuron activation. There is a one-to-one relationship between an action potential in a motoneuron and that in its muscle fibers. Within each excited muscle fiber, the action potential propagates outward from the motor end-plate and in some not entirely understood manner, triggers the contractile mechanism. The effects of this mechanism are best explained by reference to Figure 1.

Muscle behaves as if it were a visco-elastic system, the parameters of which are nonlinear functions of the muscle activity. In addition to these parametric changes, there is force generator whose output is frequently termed the "active state" [Hill, 1949]. All components are also nonlinearly dependent on muscle length and/or velocity. Many discussions of such a model have appeared in the literature [Wilkie, 1956; Ritchie and Wilkie, 1958; Bahler, et al, 1967; Vickers, 1968; Bahler, 1968.] and we shall not attempt to review them here. They have shown, with a considerable degree of success, the response of a muscle to various mechanical disturbances as well as the effects of electrical stimulation (single shock and tetanic) of the muscle or its nerve.

The total force generated by a muscle is graded by a) varying the frequency of motoneuron firing and b) varying the number of motoneurons active.

It has been long established that although firing frequency is a controlled variable of the motor system, of considerably greater significance is recruitment of varying numbers of motor units [Denny-Brown, 1928; Petajan and Philip, 1969]. Because the EMG is so closely related to the muscle's controlling neural signal we might expect it to carry much information concerning the muscle's mechanical output.

Figure 2 shows a typical EMG waveform recorded from a human soleus muscle during a rapid change in the level of voluntary isometric foot torque. Clearly, the instantaneous value of the EMG carries no more information to the eye than does a modulated radio signal. We would like some way of demodulating that EMG signal to find the mechanical response, other than using another muscle of course.

One such way might be to take the Average value of the full wave Rectified EMG, AREMG. Theoretical studies which assume that muscle force is graded solely by recruitment of motor units firing at a constant frequency with a random phase relationship to each other predict that the amplitude of the AREMG should vary with the square root of the number of motor units recruited, i.e. with the square root of muscle force [Moore, 1967; Person and Libkind, 1967; Libkind, 1968]. On the other hand, there is abundant evidence that, under isometric conditions, the AREMG varies linearly with muscle force [Lippold, 1952; Liberson, et.al., 1962; Tardieu, et.al., 1963]. In order to adjust theory to fact, it has been suggested that as muscle force increases, motor units may become progressively more synchronized in their discharge pattern resulting in a more rapid rise in AREMG and this has been demonstrated [Person and Kudina, 1968].

Other studies of EMGs and AREMGs have also shown linear relations between the electrical and mechanical events under conditions of slow

rates of contraction or of constant force, constant velocity movements [Inman et.al., 1952; Bigland and Lippold, 1954]. Nevertheless, the circumstances under which we can identify a functional relationship between the two phenomena remain extremely limited [Basmajian, 1967].

Experimental Methods

Our methods for obtaining foot torque (FT) and EMG data from the soleus muscle (SM) and the anterior tibial muscle (ATM) during isometric pursuit tracking of a visual target has been described at length previously and will be omitted here [Agarwal and Gottlieb, 1969; Gottlieb and Agarwal, 1970]. The only modification of these techniques was the added facility to control the velocity of the pursued target so that ramp as well as step efforts could be observed.

For our model simulations, a temporary hybrid computer facility was used. The muscle model was realized on an EAI TR-20 analog computer. The inputs to this model were actual EMG recordings which were output from tape by an IBM 1800 digital computer. The model response, i.e. simulated foot torque, was recorded by the digital computer and plotted by the same program used for the plotting of the physiological data.

Figure 3 shows some average tracking responses [Gottlieb and Agarwal, 1970] to different velocity ramps. Although such data are most tempting to the prospective modeler, the experiment is somewhat more complex than we would like since it involves two input signals (the soleus and anterior tibial EMGs) but only the summed output signal FT.

To simplify this, an effort involving only a single muscle group was studied by instructing the subject to actively contract his soleus muscle when tracking the target from a relaxed to a plantarflexed state but to relax rather than actively dorsiflex in following the target from

the plantar to relaxed position (observe ATM activity in Agarwal and Gottlieb, 1969, Figure 4, 5 when these instructions are not given). Thus, tracking was performed by successive contractions and relaxations of the soleus muscle and, as the subsequent data will show, the anterior tibial muscle remained quiescent. Furthermore, in order to avoid the many nonlinearities associated with changing muscle lengths, all experiments were performed under isometric conditions.

Results

One indication of the linear nature of the mechanical foot response, in both plantar flexion and dorsal relaxation is the plot of step amplitude versus peak step velocity shown in Figure 4.

The model transfer function is given by

$$\frac{FT}{EMG} = K \cdot G_1(S) \cdot G_2(S)$$

$$G_I(S) = \frac{1}{T_I S+1}$$

Separation of the transfer function like this allows us to tentatively designate $G_1(S)$ as the active state component while $G_2(S)$ may be considered the linearized mechanical component of the muscle model shown in Figure 1. The unique feature of this model is that its input is not a mechanical disturbance or electrical stimulus but recorded physiological EMG data.

Figure 5 shows the subject's and model's responses for three plantar-flexions. The model parameters are

T_1 = 0.105 sec

T_2 = 0.105 sec

K = 0.25-0.35*

*The gain of the model could not be kept constant but required small (<10%) changes from record to record with a slight overall decline over the course of an experimental session. See Tardieu et.al., 1963 for discussion of such changes during the course of an experiment and on a day to day basis.

Figure 6 shows one more step response with the addition of the output of the active state stage G_1.

The excellence with which the model reproduces the subject's plantarflexion behavior is not extended to the case of passive relaxation. Figure 7 shows one such example and the model's response is far too slow.

By trial and error, it was found that by merely reducing T_1 to

$$T_1 = 0.010 \text{ sec}$$

a reasonably accurate simulation of the rate of relaxation was obtained. Figure 8A shows one such response. Figure 8B shows another but in this (and only this) case, the input was an average EMG waveform.

In Figure 9 are shown the impulse response of the model showing the outputs of both stages. A typical twitch from a single supramaximal electrical shock of the soleus muscle motor fibers, the closest physiological equivalent to an impulse response we can obtain, is also shown for comparison.

Discussion

The strongest conclusion to be drawn from the above data is that there is a definite quantitative relationship between the EMG and foot torque during voluntary plantarflexion. This relationship is described, to a remarkably good degree, by a simple linear second order system with constant coefficients. The errors of the model that are visible in Figures 5 and 6 are attributable to the facts that the parameters are not constant, even under isometric conditions, and that our EMG recordings are from only one of the three muscles (the others being the two heads of the gastrocnemius) that produce the measured torque and have different dynamics (the gastrocnemius being a fast muscle as compared to soleus).

In calling the output of stage G_1 the active state, our first justification is in the structure of the muscle model (Figure 1) itself. This variable is in no way directly measurable, intrinsic to the muscle tissue. It represents a hypothetical quantity that may be regarded as input by the muscle's contractile machinery to its noncontractile elements. It is only after passing through this mechanical filter that we may measure the resultant force. Indirect measurements of the active state have only been made on animal (usually isolated muscle) preparations with electrical stimulation [Hill, 1949; Ritchie and Wilkie, 1955; Bahler et.al., 1967]. Such measurements have shown that in response to a single shock to the muscle or its motor nerve, the active state rises rapidly to its maximum and then decays more slowly.

These results are in qualitative agreement with the model response in Figure 9. Such agreement also exists between the model's torque impulse response and the subject's twitch response. The physiological data differs from the model in the apparent clipping of the initial velocity peak and the notch in the torque at the end of the twitch (corresponding to the second negative peak in the acceleration curve). This notch is the "myotatic hump" [Creed, et.al., 1932] identified as a product of the stretch reflex loop which is always present in our experiments. Therefore, our muscle's "impulse response" is obscured, in part, by the feedback loop in which it is embedded but this only affects the last part of the twitch curve.

When we try to apply the same model to simple relaxation we find, as is obvious from Figure 7, that the model's response is far too slow. By trial and error we found that reducing the time constant of one of the stages by a factor of ten would considerably improve the model response

in one sense. That is, the model's rate of fall in response to the EMG step is adequate but that fall begins from 50 to 100 msec to soon. Furthermore, the torque is far too sensitive to brief EMG bursts that often occur after the initial turning off of the EMG.*

In considering the fact that relaxation is more rapid than contraction** we may note two other types of evidence. First, the active state falls far more slowly than it rises and the gross mechanical dynamics of relaxed muscle are slower than those of active muscle [Wilkie, 1950; Okabe et.al., 1967; Agarwal et.al. 1970] and these facts would suggest a slower relaxation rate. Yet second, and more in agreement with our results are those obtained from stimulation of motor fibers by frequency modulated (sinusoidal) pulse trains [Partridge, 1966]. Those experiments also show relaxation to be more rapid than contraction. The resolution of these conflicts is not now within the scope of our knowledge.

We return finally to the basic question of what information is actually contained in the EMG. If it is truly a measure of motoneuron activity it then represents the muscle's controlling signal and in describing the relationship between EMG and force, we are describing the fundamental dynamics of muscle activity under normal operating conditions. If, however, the EMG is some function of motoneuron activity involving considerable dynamic processing of its own, then the relationship between EMG and force tells us far less about muscle dynamics and we must regard the success of our model as more fortuitous than substantial. At this stage however, such success is encouraging and may prove most useful.

---

* The variability in model torque prior to relaxation is not significant. The short time constant that allows this is only appropriate to the relaxation process and not to the torque maintenance which precedes it.

** In absolute terms, relaxation is slightly slower than contraction but the driving EMG input for contraction permits the model to be strongly over-driven in that direction. The model cannot be over-driven to relaxation since a negative EMG (for a single muscle simulation) is not permitted.

REFERENCES

1. Agarwal, G. C., B. M. Berman , and L. Stark. "Studies in postural control system. I. torque disturbance input." IEEE Trans. Syst. Science and Cybernetics, 1970. (in press)

2. Agarwal, G. C. and G. L. Gottlieb. "Analysis of step tracking in normal human subjects." IEEE Trans. on Man-Machine Syst., December, 1969.

3. Bahler, A. S. "Modeling of mammalian skeletal muscle." IEEE Trans. Bio-Med. Eng., 1968, 15:249.

4. Bahler, A. S., J. T. Fales, K. L. Zierler. "The active state of mammalian skeletal muscle." J. Gen. Physiol., 1967, 50:2239.

5. Basmajian, J. V. Muscles Alive, 2nd Edition, The Williams & Wilkins Company, Baltimore, 1967, 421 p.

6. Bigland, B. and O. C. J. Lippold. "The relation between force, velocity and integrated electrical activity in human muscles." J. Physiol. 1954, 123:214.

7. Creed, R. S., D. Denny-Brown, J. C. Eccles, E. G. T. Liddell and C. S. Sherrington. Reflex Activity of the Spinal Cord, Oxford University Press, London, 1932, 183 p.

8. Denny-Brown, D. "On the nature of postural reflexes." Proc. Roy. Soc. (London) Ser. B. 1928-29, 104:252.

9. Gottlieb, G. L. and G. C. Agarwal. "Interactions between the voluntary and postural mechanisms of the human motor system" J. Neurophysiol. (accepted for publication), 1970.

10. Hill, A. V. "The abrupt transition from rest to activity in muscle." Proc. Roy. Soc. (London), Ser. B, 1949, 136:399.

11. Inman, B. T., H. J. Ralston, J. Saunders, B. Feinstein and E. W. Wright, Jr. "Relation of human electromyogram to muscular tension" Electroenceph. Clin. Neurophysiol. 1952, 4:187.

12. Liberson, W. T., M. Dondey, and M. M. Asa. "Brief repeated isometric maximal exercises." Am. J. Phys. Med. 1962, 41:3.

13. Libkind, M. S. "II. Modelling of interference bioelectrical activity." Biophysics (Biofizika) 1968, 13:811 (English translation, Pergamon Press).

14. Lippold, O. C. J. "The relation between integrated action potentials in a human muscle and its isometric tension." J. Physiol. 1952. 117:492.

15. Moore, A. D. "Synthesized EMG waves and their implications." Am. J. Phys. Med., 1967, 46:1302.

16. Okabe, Y., H. E. Rhodes, L. Stark, and P. A. Willis. "Transient response of human motor coordination system." MIT Res. Lab. Elect. 1962, QPR 66:389.

17. Partridge, L. D. "Signal-handling characteristics of load-moving skeletal muscle." Amer. J. Physiol. 1966, 210:1178.

18. Person, R. S. and L. P. Kudina. "Cross-correlation of electromyograms showing interference patterns." Electroenceph. Clin. Neurophysiol. 1968, 25:58.

19. Person, R. S. and M. S. Libkind. "Modelling of interference bioelectrical activity." Biophysics (Biofizika) 1967, 12:145 (English translation, Pergamon Press).

20. Petajan, J. H. and B. A. Philip. "Frequency control of motor unit action potentials." Electroenceph. Clin. Neurophysiol. 1969, 27:66.

21. Ritchie, J. M. and D. R. Wilkie. "The effect of previous stimulation on the active state of muscle." J. Physiol., 1955, 130:488.

22. Ritchie, J. M. and D. R. Wilkie. "The dynamics of muscular contraction." J. Physiol., 1958, 143:104.

23. Tardieu, G., C. Tardieu, C. Monfraix, L. Gagnard, and J. Velin. "Etude critique de l'electromyographic comme methode d'evaluation des infirmites motrices cerebrales." Rev. Neurogique, Paris 1963, 108:87.

24. Vickers, W. H. "A physiologically based model of neuromuscular system dynamics." IEEE Trans. on Man-Machine Systems, 1968.

25. Wilkie, D. R. "The relation between force and velocity in human muscle." J. Physiol. 1950, 110:249.

26. Wilkie, D. R. "The mechanical properties of muscle." Brit. Med. Bull., 1956, 12:177.

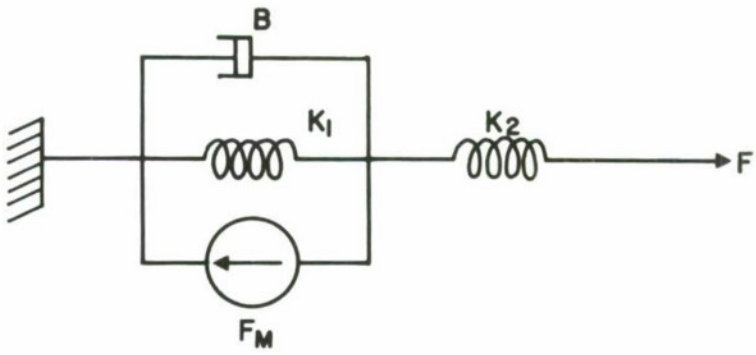


Figure 1. A mechanical model of skeletal muscle. The series and parallel elastic elements (K_2 and K_1) and the viscosity B are defined only by the mechanical behavior of the muscle and should not imply a mechanism. The force generator Fm represents the contractile "active state" of the muscle.

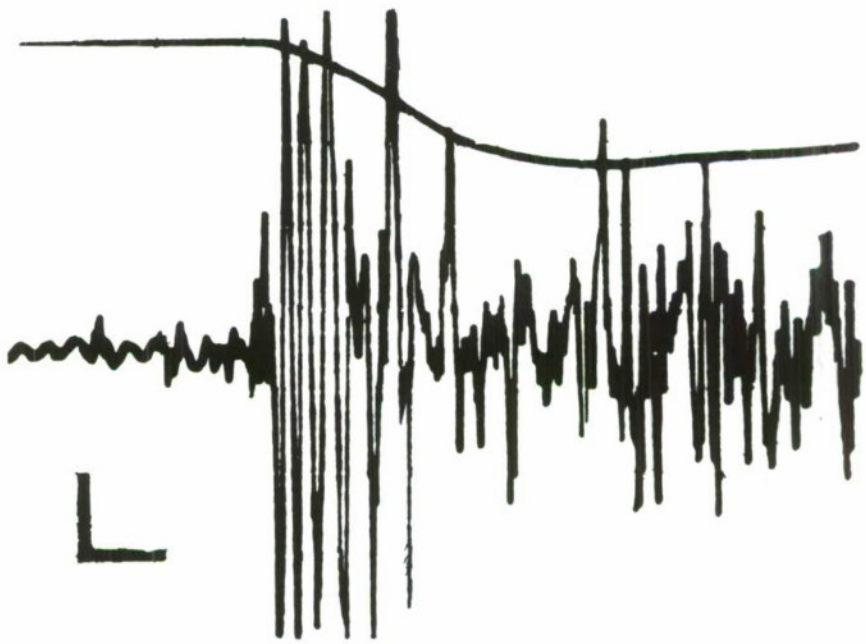


Figure 2. Foot torque (top trace) and soleus muscle EMG during a rapid step effort. Scales: Foot torque 0.5 Kg.m./unit, EMG 0.1 mv/unit, time .05 sec/unit.

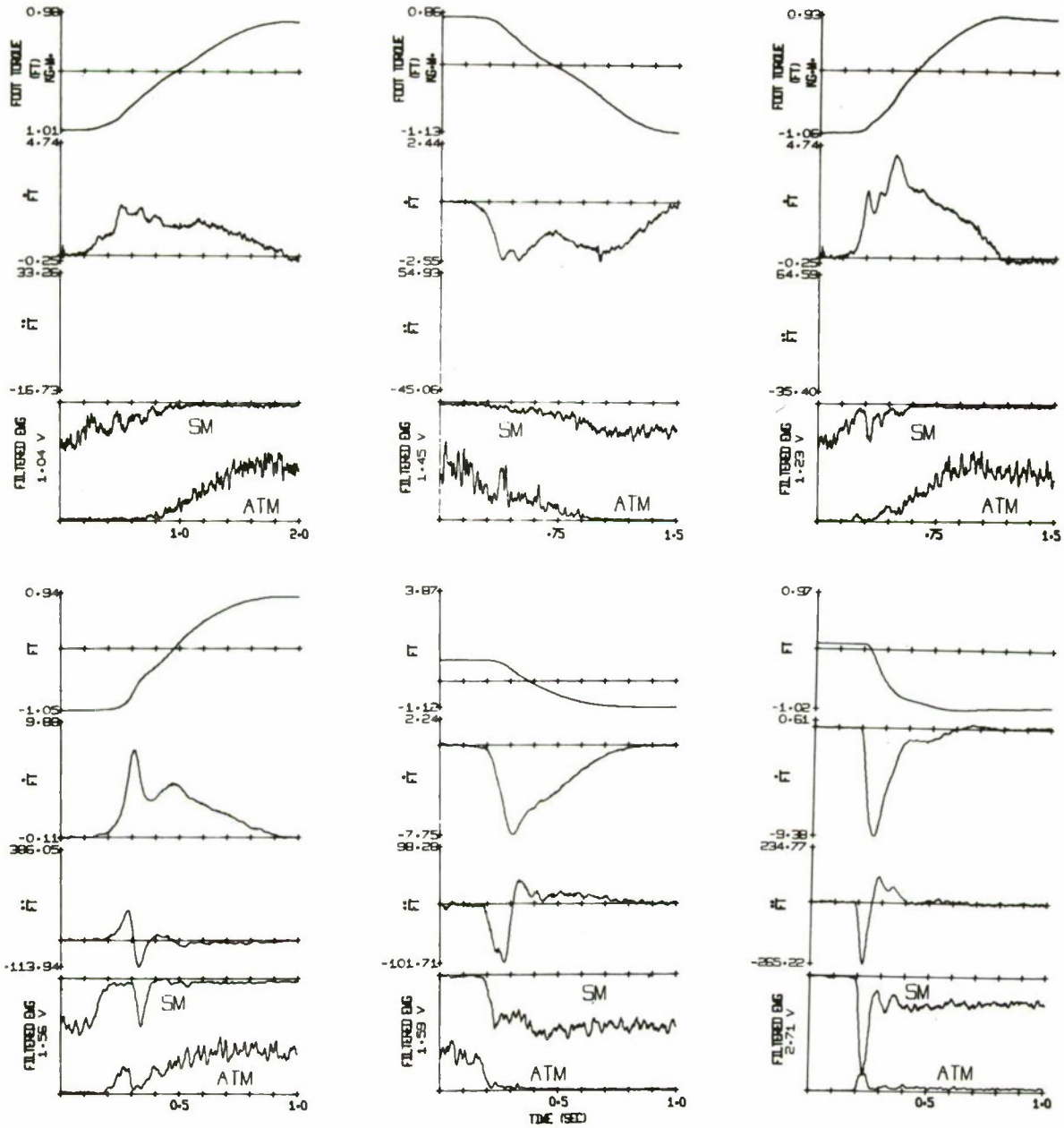


Figure 3. Recorded average foot torque (FT) and the first two computed time derivatives ($\dot{FT}$, $\ddot{FT}$) in tracking different velocity targets. Recorded EMG's are from the soleus muscle (SM-increasing downward) and the anterior tibial muscle (ATM-increasing upward). Each average is of sixteen responses. Target end points are $\pm$ 1 Kg.m. (except the last which are 0 Kg.m., -1 Kg.m.), negative torque is in the plantar direction. Target velocities, from left to right, top to bottom are 1.33, 2.0, 2.67, 4.0, 5.75 Kg.m./sec. and a step response. Target movement begins 0.2 sec. after the start of the recording for the five ramp experiments. The plot of $\dot{FT}$ is sometimes omitted by the computer if it is too noisy.

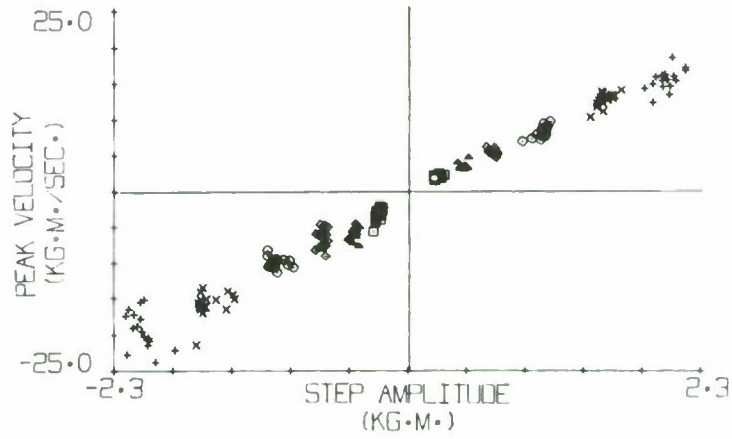


Figure 4. A plot of the peak tracking velocity achieved during step tracking versus the amplitude of the target steps. The third quadrant points are of active plantar-flexion. The first quadrant points are of relaxation from an initial plantarflexion to a neutral rest torque.

For the first quadrant, the correlation coefficient is r = 0.989 and the regression curve is V = 7.84A+ 0.16, (112 points). For the third quadrant, the correlation coefficient is r = 0.957 and the regression curve is V = 8.73A-1.32, (113 points).

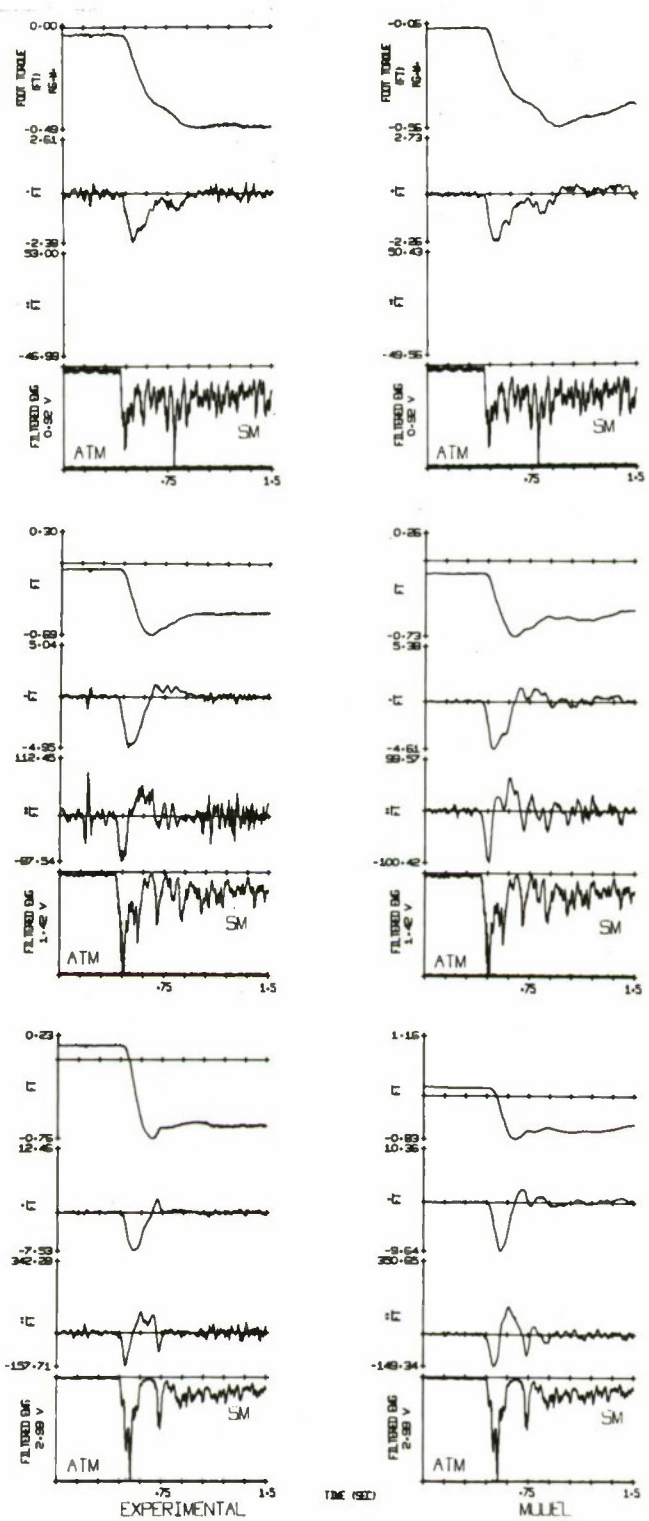


Figure 5. LEFT: Step torque responses and associated EMG's (soleus EMG increasing downward). RIGHT: Model torque responses to the recorded EMG data. These examples illustrate responses to different amplitude steps, with and without overshoot by subject.

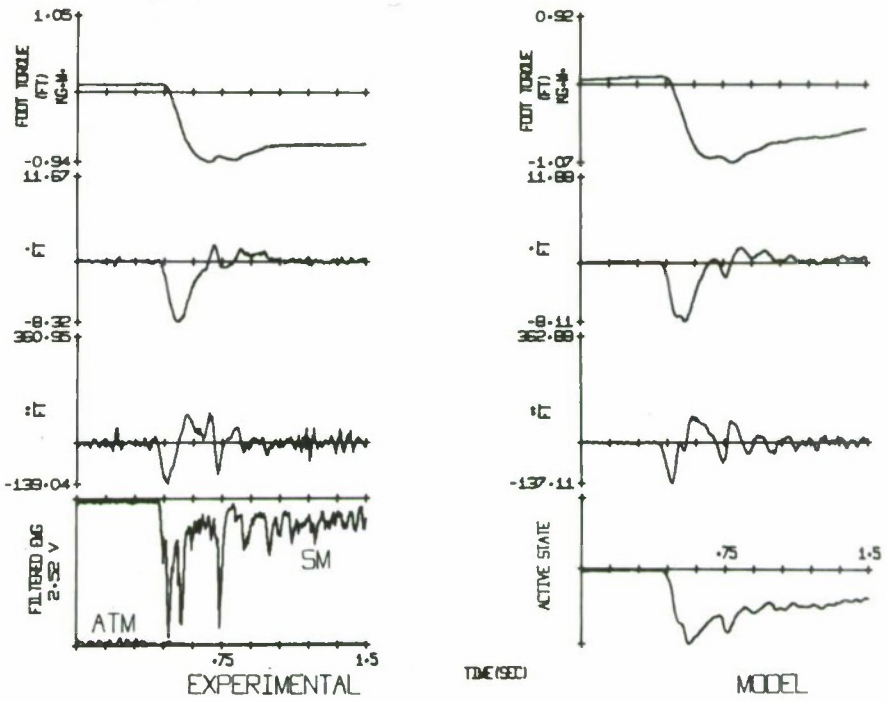


Figure 6. LEFT: Subject's step torque response and EMG. RIGHT: Model response to subject's EMG. Instead of replotting EMG on the right, the output of the active state stage G_1 is the fourth curve plotted.

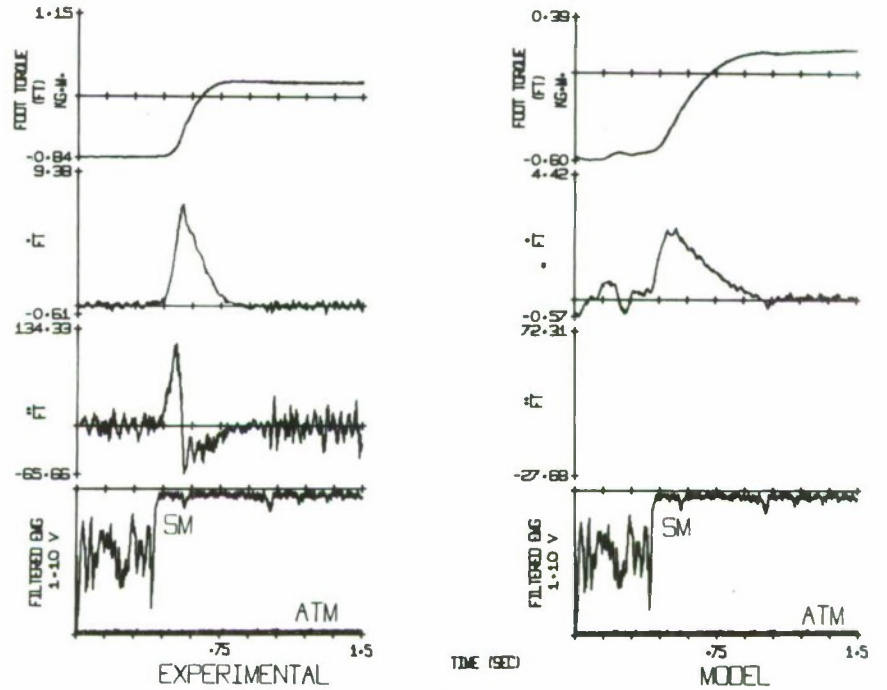


Figure 7. Relaxation of foot torque and the model response. Model time constants are those used in Figures 5 and 6.

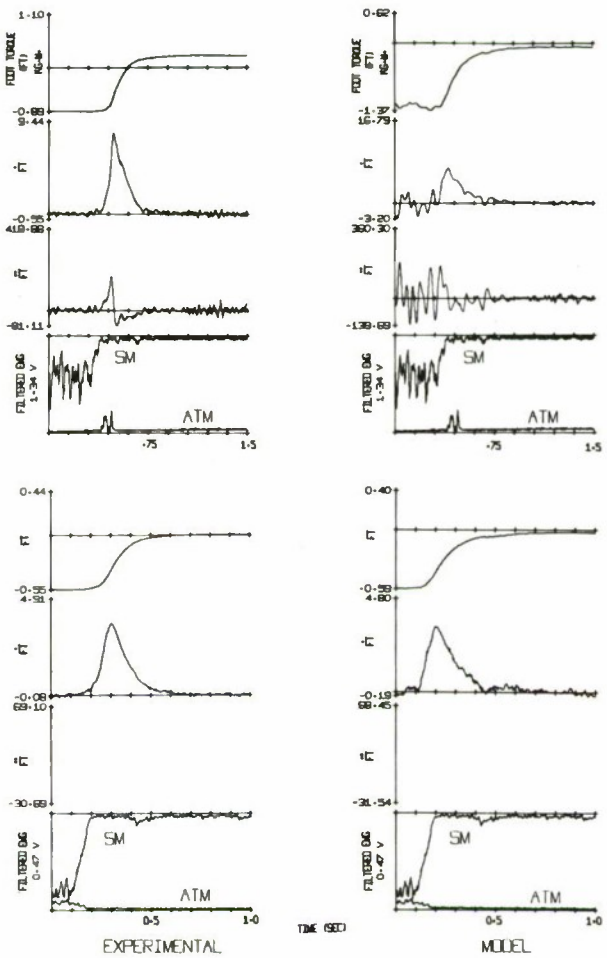


Figure 8. TOP: Experimental and model relaxation response. Model time constant reduced in one stage. BOTTOM: Experimental average relaxation response and model (with reduced time constant) response. Average used to reduce EMG variance. Note that model response, although of approximately correct time constant, begins 50-80 msec. earlier than the experimental data.

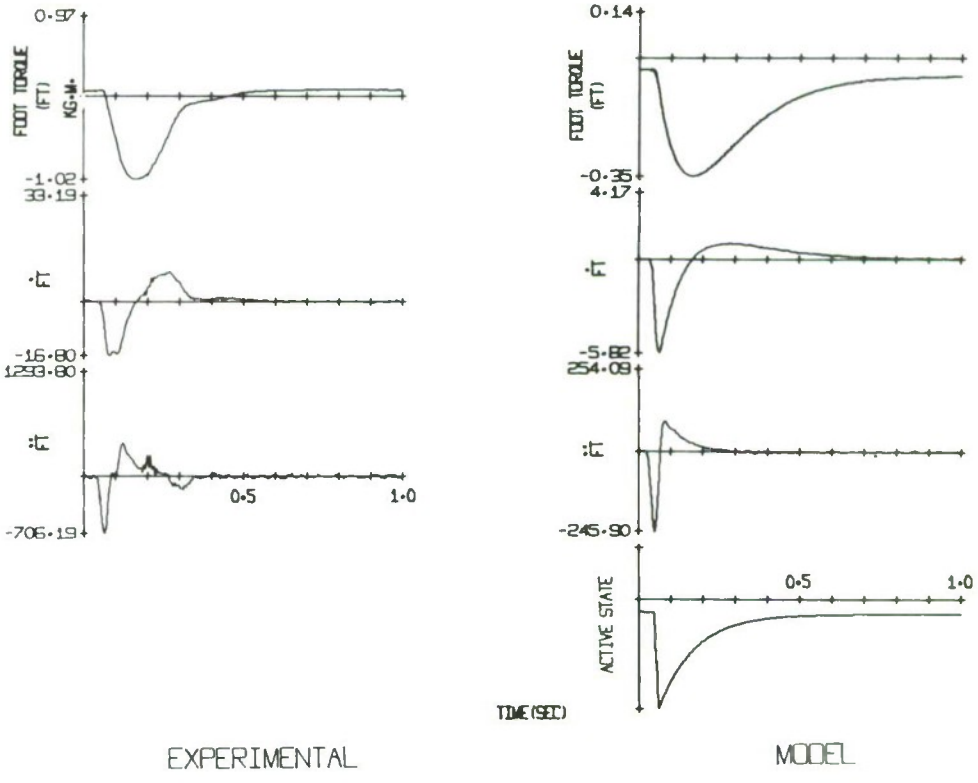


Figure 9. LEFT: Foot twitch in response to a single supramaximal shock to the tibial nerve. RIGHT: Model impulse response and the output of the active state stage G_1.

ON THE GRADING OF TENSION DURING VOLUNTARY CONTRACTION OF SKELETAL MUSCLE

L. Ostroy, A. V. Phatak, and G. A. Bekey

Abstract

The gradation of tension during voluntary contraction of skeletal muscle is due to two factors: (a) the firing statistics of individual motor units and (b) the temporal and spatial order in which motor units are activated. It is well known from experimental data [1] that the standard deviation σ of the surface EMG potential and the average tension T of the muscle are linearly related, i.e. $\sigma_{EMG}=kT$. The purpose of this study is to investigate the inverse problem of determining possible firing statistics and recruitment order of single motor units which yield the above EMG-tension relationship.

It is sometimes assumed that the level of tension in skeletal muscle is set by neuron firing rate and random recruitment of motor units. This hypothesis was simulated on a digital computer and a synthetic EMG was produced. It was found that the tension was proportional to the variance of the EMG signal, rather than its standard deviation (as shown experimentally).

Recent physiological data suggest the following alternative hypotheses:

(1) Recruitment of motor units occurs at the center of the muscle at low tension levels and proceeds towards the surface as tension level increases.

(2) Recruitment of motor units is in order of size.

Preliminary analysis of these hypotheses indicates that they give results consistent with experimental data.

I. Introduction

This paper presents a mathematical model for skeletal muscle under isometric contraction based on the properties of single motor units. It gives an analysis of a plausible mechanism that relates the two muscle macro-outputs of isometric tension and surface myoelectric activity (EMG) to the single neuronal input. The model assumes that the nervous input to the muscle is a stochastic process, thereby implying the stochastic nature of muscle tension and surface EMG. Any correlation between the statistics of the two muscle output processes would depend on the assumptions made on the statistics of the nervous input. The latter depend on the following two factors: a) the firing statistics of individual motor units and b) the temporal and spatial order in which motor units are activated. It is well known from experimental data [1] that the standard deviation σ of the surface EMG potential and the average muscle tension T are linearly related, i.e. $\sigma_{EMG} = kT$, where k is a constant. The purpose of this study is to investigate the inverse problem of determining possible firing statistics and recruitment order of single motor units which yield a linear EMG-tension relationship. This approach is necessary because the recording of total nervous excitation in intact muscle under isometric contraction is difficult, if not impossible, at the present time [2].

The structure of the proposed model is based on the work of Coggshall [3]. The model was used to synthesize an EMG on a digital computer. The statistics of the synthetic EMG, in turn, were used to determine the validity of the model and the underlying assumptions.

II. Background and Details of the Model

As mentioned earlier, the model developed here for the whole muscle uses individual motor units as building blocks. A brief discussion of muscle

physiology relevant to this study is thus included in the following.

A muscle is composed of a large number of individual motor units. A motor unit consists of a single α motoneuron in the spinal cord, the nerve axon and its branches, and the muscle fibers which the latter innervate via end-plate synapses. The muscle fibers of a single motor unit are not necessarily contiguous and are distributed throughout the entire muscle.

When a nervous impulse travels down the nerve axon of the motor unit and strikes the end plates of the muscle fibers, two macroscopic events occur: 1) The fibers of the motor units contract, producing what is called a "twitch", denoted by f(t), as shown in Fig. 1a [4, 5]; and 2) The fibers depolarize, creating a short-lived electrical potential across the membrane of the fibers. This potential, denoted by v(t), has the form shown in Fig. 1b and is known as a motor unit potential or "mup" [6].

Experimentally it has been shown that both f(t) and v(t) are random processes; hence the shapes shown in Figs. 1a and 1b are ensemble averages.

When a muscle is contracted voluntarily, there are two processes by which tension can be increased: 1) The firing rate of the active motor units is increased or 2) Dormant motor units are recruited. In this model the nervous input to each motor unit is represented by a train of impulses, $\sum_{i=1}^{\infty} \delta(t-ti)$, and β_i is defined as the mean firing rate of the individual i-th motor units. Based on this, each motor unit can be represented by a linear model as in Fig. 2 where f(t) and v(t) refer to the typical twitch and mup responses shown in Fig. 1a, b. Physiologically, an entire muscle is the sum of a large number of motor units. If one further stipulates that the net effect of recruiting dormant motor units is to change the average firing rate β of the total nervous input, then assuming a Poisson process for the nervous activity $\beta = \sum_i \beta_i$, where β_i is as defined above. The entire muscle model may now be described as in Fig. 3. Note that because

of the assumption that the nervous input is a Poisson process, the only statistical parameter of concern is the mean firing rate β of the total nervous input.

III. Criterion for Validity of the Model

Experimentally, it has been shown that the f's and v's are random processes with mean shapes approximated in Fig. 1. The EMG and total tension outputs of the total muscle are therefore summations of the respective curves in Fig. 1 over all active motor units. It has also been shown experimentally that the standard deviation of V_{EMG} varies linearly with T and goes through the origin as shown in Fig. 4 [1]. Therefore,

$$T = k\sqrt{VAR(V_{EMG})} \tag{1}$$

where k is a constant.

In what follows the linear relationship given in Fig. 4 serves as a necessary condition that must be satisfied by the synthetic EMG for the corresponding muscle model to be valid.

IV. Critical Assumptions in Modeling

In order to proceed further, some additional simplifying assumptions have to be made about muscle physiology. First, it will be assumed that the ensemble average motor unit twitch remains the same regardless of how much tension is generated. What this means is illustrated by the following equations:

$$T = k_F \beta \tag{2}$$

where T is the average isometric tension and β = total nervous mean firing rate

$$k_F = \int_{-\infty}^{\infty} \bar{f}(t)dt \tag{3}$$

$\bar{f}(t)$ = ensemble average twitch

Secondly, it will be assumed that the percentage of active motor units which contribute to the surface EMG voltage remains constant. Hence,

$$\beta_v = a\beta \tag{4}$$

where a is a constant and β_v = mean firing rate of the active motor units contributing to the EMG.

The nervous input statistics are now hypothesized to be periodically varying, randomly timed impulses at rate

$$\beta = A + B \sin 2\pi t/T \tag{5}$$

where A, B, and T are constants.

The reason this form is chosen for β is that EMG's have been shown to exhibit periodicity and "bunching". This phenomenon is called synchrony in the literature [7]. Dealing with this form for β is mathematically difficult due to the nonstationarity of the process. A numerical calculation of EMG is therefore necessary and this leads to the computer synthesis of the EMG as described in the next section. A synthesized EMG is shown in Fig. 5.

V. EMG Synthesis Method

The method of synthesis is as follows. The computer generates a list of times from the relations

$$t_{i+i} = t_i + \Delta t \tag{6}$$

$$\Delta t = -\frac{1}{\beta_v} \log (\text{RANDM(IRAND)}) \tag{7}$$

where RANDM(IRAND) generates random numbers from 0 to 1 in a uniform distribution. A motor unit potential and a twitch are assumed to begin almost simultaneously at each one of these times. The sum of all motor unit potentials is then made. See Fig. 6. This sum is the synthesized EMG of Fig. 5. A whole

family of EMG's was synthesized using this method and a curve of $(\sqrt{VAR(V_{EMG})_{model}})$ vs. β_v both averaged over 25 periods was obtained and the relation between them is shown in Fig. 7. The curve is almost a perfect parabola with very high confidence. Therefore,

$$VAR(V_{EMG})_{model} = k\beta_v. \qquad (8)$$

Since the tension T is proportional to β_v, Fig. 7 and equation 8 indicate that the model does not give the correct result as experimentally obtained and shown in Fig. 4.

VI. Reexamination of Assumptions and Alternate Assumptions

An examination of the assumptions which led to this result is therefore in order. The first major assumption was that the ensemble average twitch is constant. If, however, we hypothesize recruitment of motor units with larger twitches first, for example, as shown in Fig. 8, and also assume that

$$\int_{-\infty}^{\infty} \bar{f}(t)dt = \frac{C}{\sqrt{\beta}}$$

then the correct result would be obtained in the final analysis, i.e., this assumption leads to the relation

$$T = C(\sqrt{VAR(V_{EMG})_{model}})$$

The question to be answered is whether motor units are recruited in order of twitch size and if so, what is the relation between size and recruitment order.

The second major assumption was that the same constant percentage of active motor units contribute to the EMG signal for all values of the tension. If, however, recruitment is hypothesized to take place as in Fig. 9, from the belly to the surface of the muscle and if it is also postulated that the recruitment

has a patterned spatial distribution such that

$$\beta_v = a\beta \text{ as before}$$

$$\text{but } a = a'\beta \qquad (9)$$

where a' is a new constant

then again the final analysis would have the correct result as follows:

$$T = k_F\beta \qquad (2)$$

from (4), (9) $$\beta_v = a'\beta^2 \qquad (10)$$

$$T = \frac{k_F}{\sqrt{a'}}\sqrt{\beta_v} \qquad (11)$$

from (8), (11) $$T = \frac{k_F}{\sqrt{a'k}}\left(\sqrt{VAR(V_{EMG})_{model}}\right) \qquad (12)$$

which is a straight line through the origin as required. There exists qualitative data supporting organized spatial recruitment such as that postulated here [8].

VII. Conclusion

The conclusions drawn from this study are mostly negative. In formulating the model, a great many factors had to be hypothesized. It is therefore impossible to judge whether or not the model can indeed represent the actual skeletal muscle under isometric contraction that it was intended to represent. The model does, however, give a very clear view of what the unknown factors are and it does suggest several areas for further experimental investigation. These areas are:

1. Size of motor units versus order of recruitment.
2. Spatial recruitment versus order of recruitment.
3. Statistical parameters of the nervous input process. For example,
   a) correlations (auto and cross-) b) firing rates
4. Electrical propagation properties in muscle to determine
   a) frequency dispersion characteristics
   b) velocity of propagation
   c) attenuation

REFERENCES

1. Lippold, O. C. J., "The Relation Between Integrated Action Potentials in a Human Muscle and Its Isometric Tensions", Journal of Physiology, Vol.117, 1952, pp. 492-499.

2. Moore, A. D., "Synthesized EMG Waves and Their Implications", American Journal of Physical Medicine, Vol.16, No. 3, 1967, pp. 1302-16.

3. Coggshall, John C., "Mathematical Models of Muscle", USCEE Report 303, University of Southern California, August 1968.

4. Gordon, G. and H. S. Holbourn, "The Mechanical Activity of Single Motor Units in Reflex Contractions of Skeletal Muscle", Journal of Physiology, Vol. 110, 1949, pp. 26-35.

5. Wuerker, R. B., A. M. McPhedran, and E. Henneman, "Properties of Motor Units in a Heterogeneous Pale Muscle (M. Gastrocnemius) of the Cat", Journal of Neurophysiology, Vol.28, 1965, pp. 85-99.

6. Sacco, G., F. Buchthal, and P. Rosenfalck, "Motor Unit Potentials at Different Ages", Archives of Neurology, Vol. 6, 1962, pp. 366-373.

7. Lippold, O. C. J., J. W. T. Redfearn and L. Vuco, "The Rhythmical Activity of Groups of Motor Units on the Voluntary Contraction of Muscle", Journal of Physiology, Vol. 137, 1957, pp. 473-387.

8. Clamann, Hans, "A Quantitative Analysis of the Firing Pattern of Single Motor Units of a Skeletal Muscle of Man, and Their Utilization in Isometric Contraction", Ph.D. Dissertation, The Johns Hopkins University, 1968.

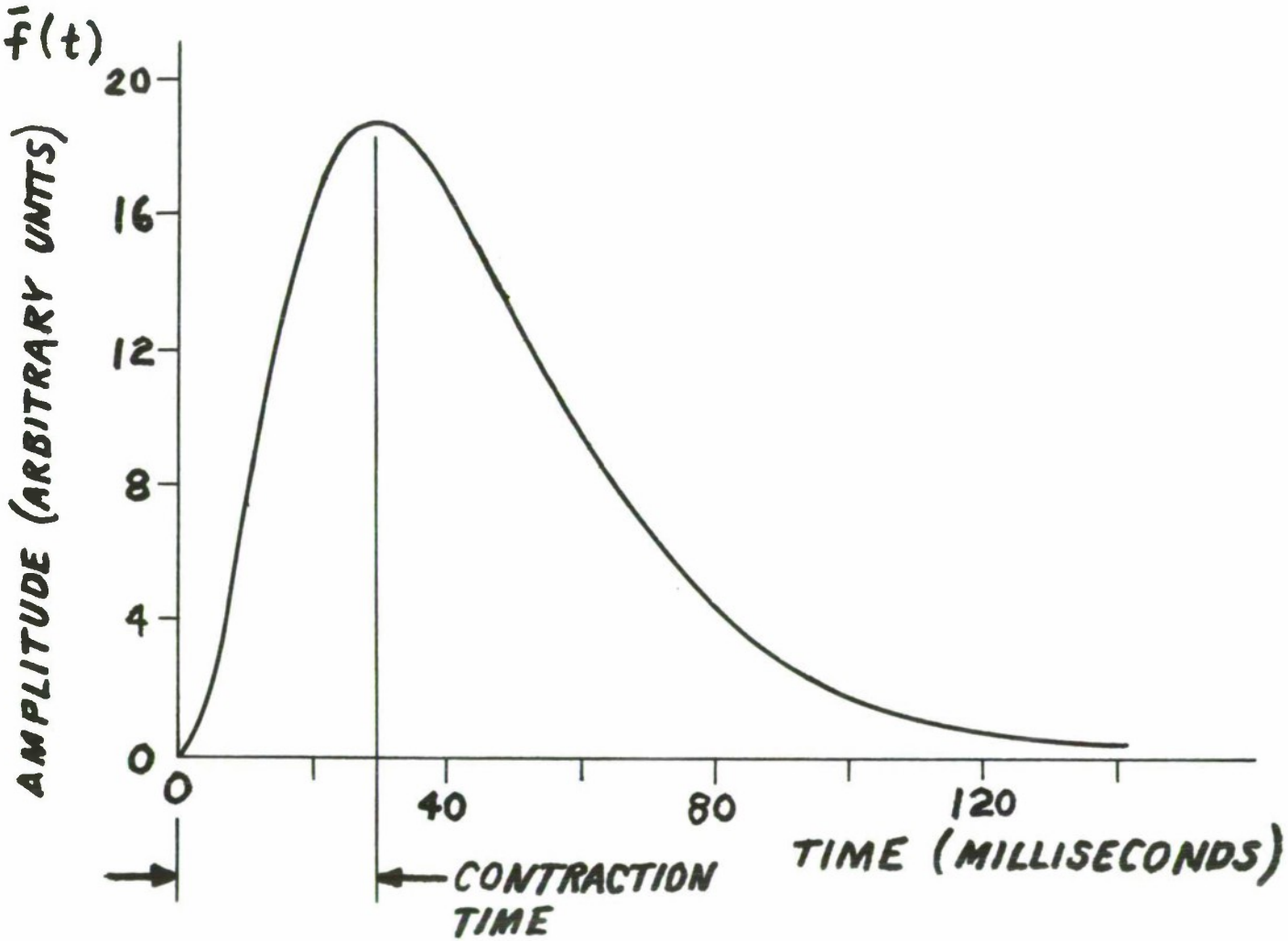


FIG. 1A - A TYPICAL MOTOR UNIT TWITCH (FROM A CAT "FAST" MUSCLE)

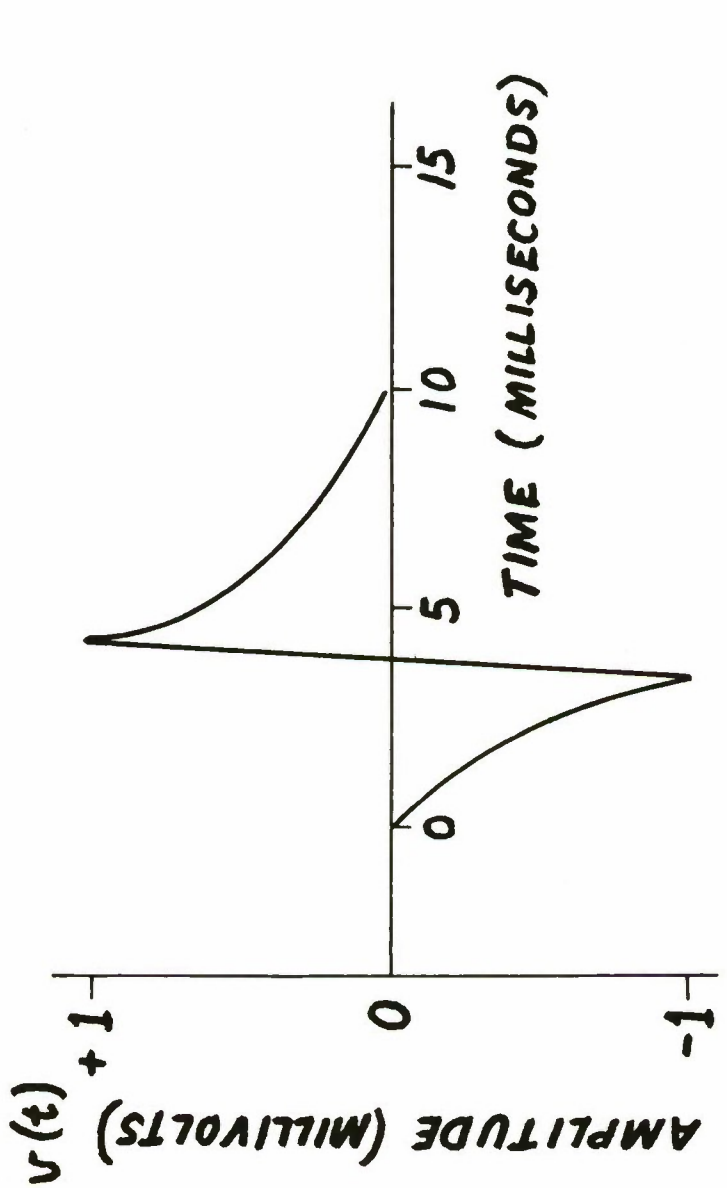


FIG. 1B – A TYPICAL MOTOR UNIT POTENTIAL

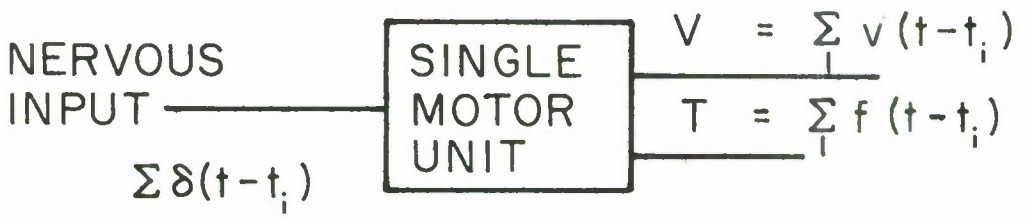


FIG.-2.-Model of a single motor unit

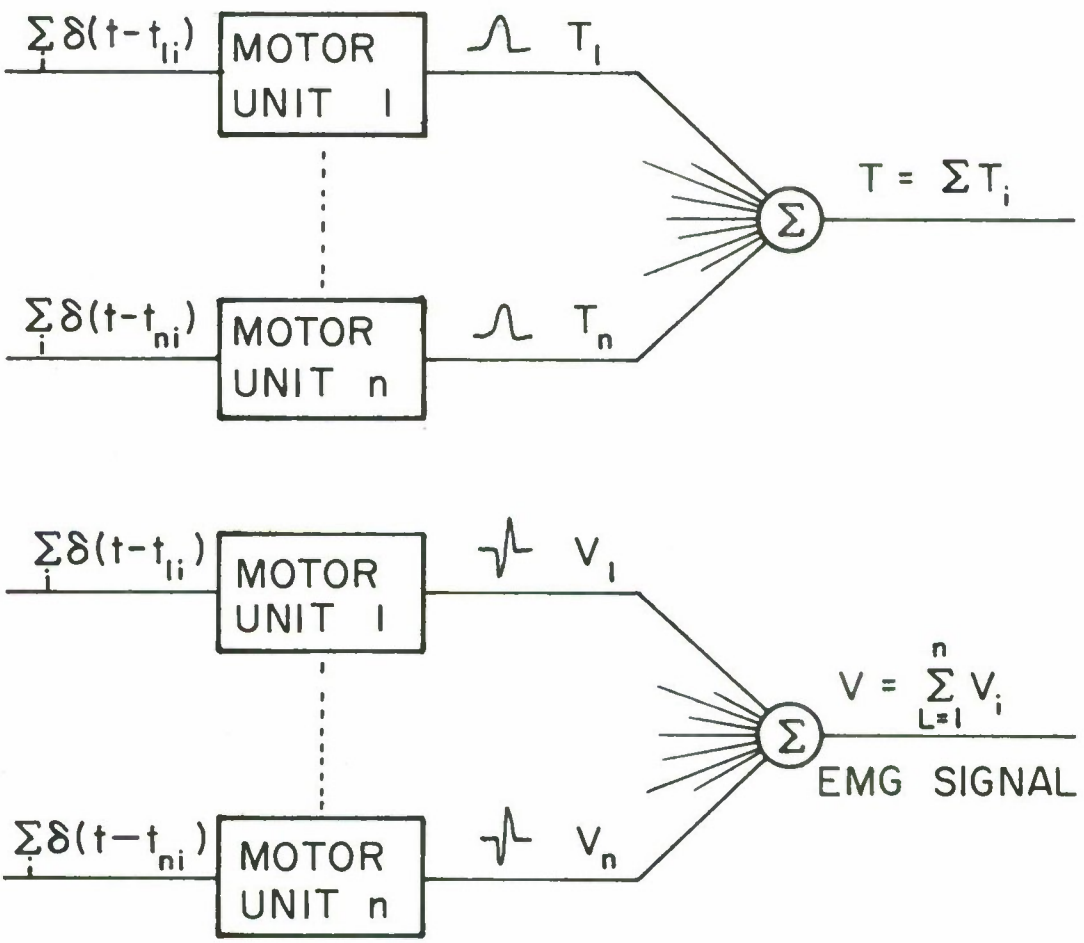


FIG. 3 - ENTIRE MUSCLE MODEL

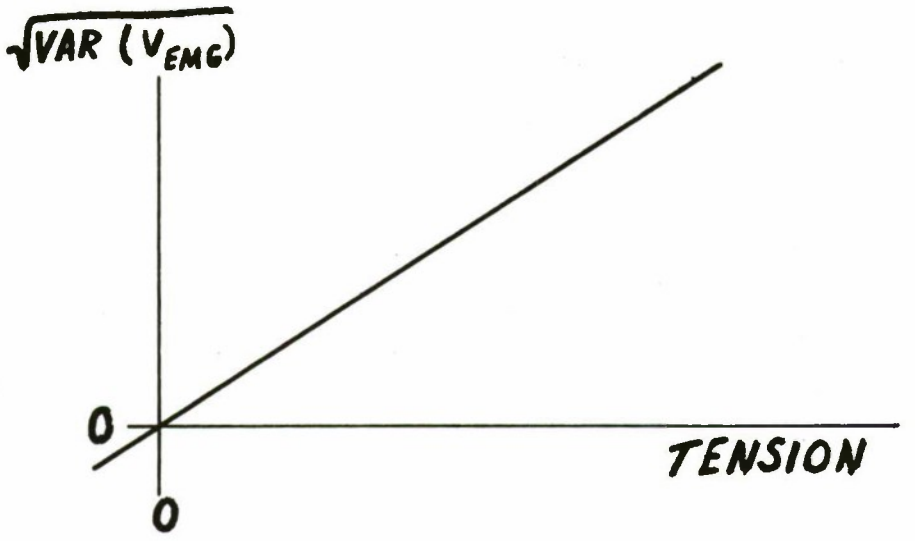


FIG. 4 – EXPERIMENTAL RELATION BETWEEN TENSION & EMG

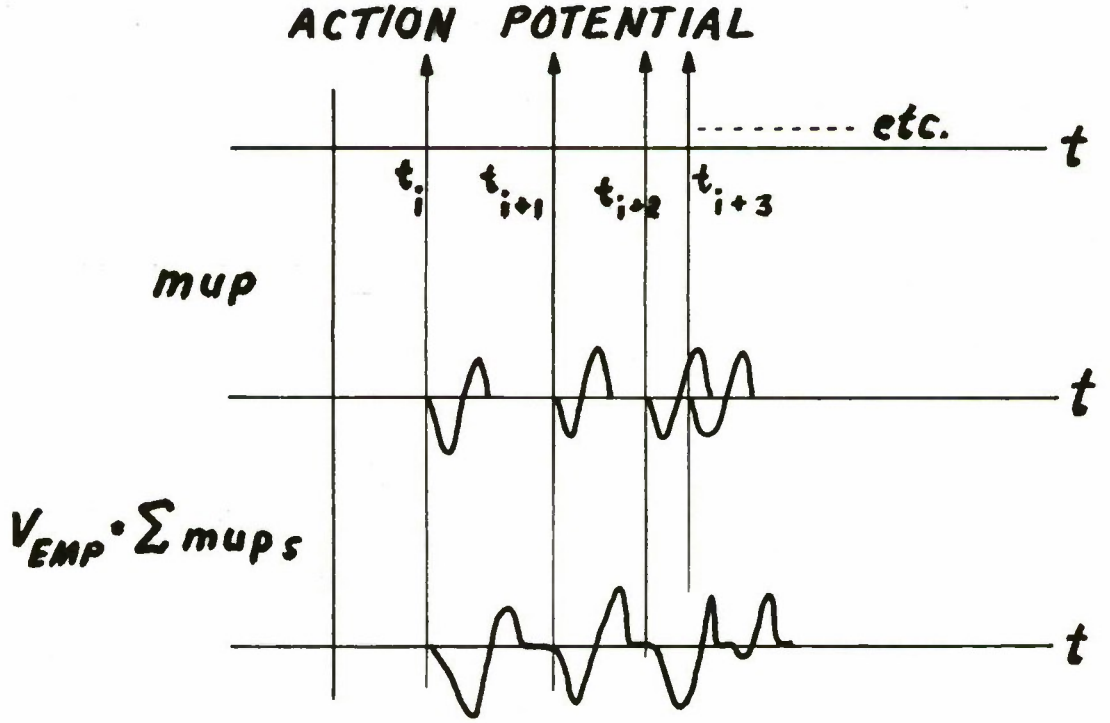


FIG. 6 – EMG SYNTHESIS FROM SUMMATION OF MOTOR UNIT POTENTIALS

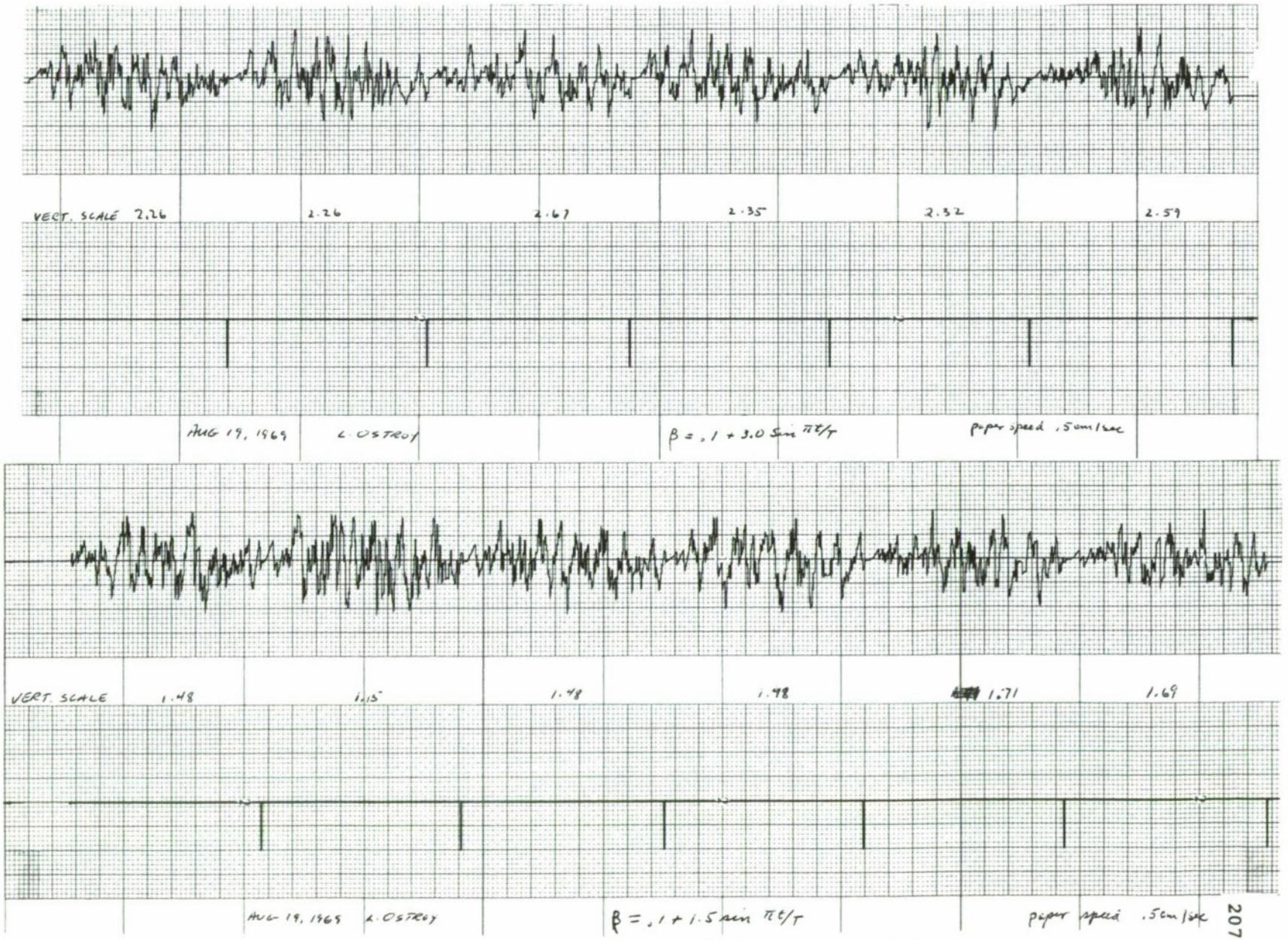


FIG. 5 - SIMULATED ELECTROMYOGRAMS

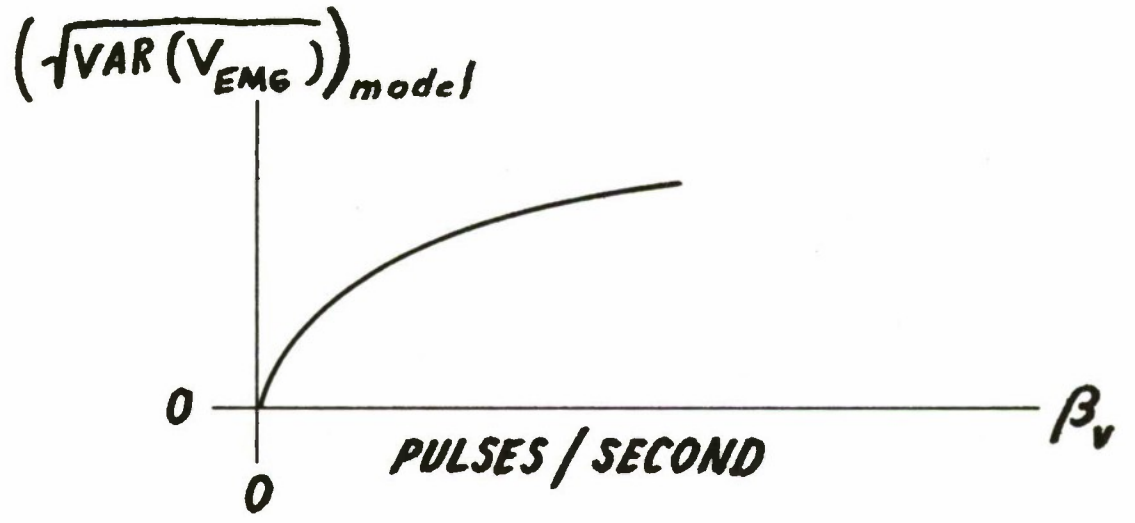


FIG. 7 - RELATION BETWEEN β_v & VAR (V_{EMG}) OF SYNTHESIZED EMG

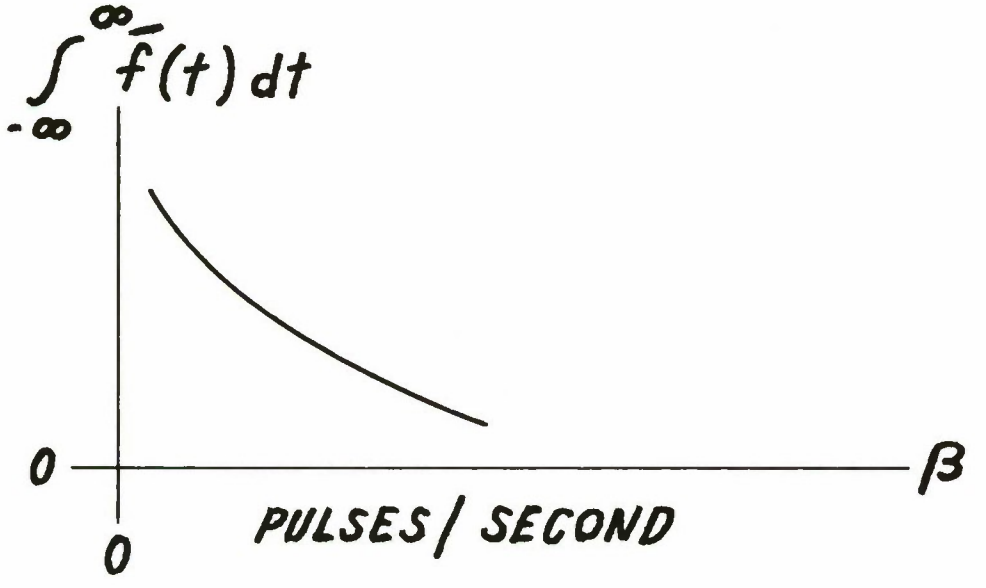


FIG. 8 - POSSIBLE RELATION BETWEEN TWITCH SIZE & RECRUITMENT ORDER

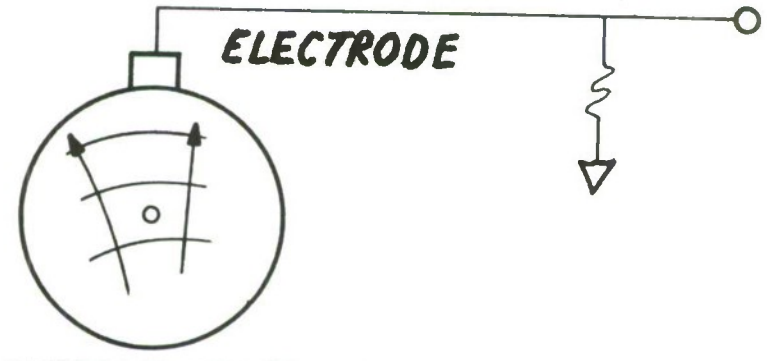


FIG. 9 - RECRUITMENT BY SPATIAL PATTERN

THE DYNAMIC CHARACTERISTICS OF HUMAN SKELETAL MUSCLE MODELLED FROM SURFACE STIMULATION†

J. A. Tennant
Stanford University††

ABSTRACT

A mathematical model for the behavior of the muscle group comprised by the biceps and the brachialis was sought taking into account the changes in muscle tension due to the inertia of the moving masses.

The approach involved the postulation of a mechanism composed of a Contractile and a Series Elastic Element. Isotonic and isometric experiments were formulated to allow the explicit characterization of the equations for the postulated mechanism. These experiments involved the artificial stimulation of the muscles via a surface electrode, located at the motor point of the muscle group. In the isotonic experiments, external loads were simultaneously applied to the forearm by means of an electric torque motor. The electric motor was controlled not only as a torque producing device, but also as a regulator to compensate for developing inertial loads during accelerated motions, as well as for the changing geometry of the arm/muscle system. The stimulus waveforms were provided digitally.

INTRODUCTION

The hypothesis that a physical system comprised of a contractile and a series elastic element can be used to describe the behavior of human muscle is tested, and a mathematical model for the dynamic characteristics of a muscle group (acting about the elbow and comprised of the biceps and the brachialis) is sought. Experimental investigations of the dynamic properties of muscles have been conducted for several decades. The most notable results are those obtained for frog and cat muscles in vitro and in vivo. Most of these experiments considered an isotonic shortening muscle to characterize its contractile element by applying a constant external load to the muscle throughout its contraction without taking into account the changes in muscle tension due to the inertia of the moving masses. This study investigates the dynamic characteristics of human skeletal muscle in situ, taking into consideration the inertia of the system, and compares the results obtained with those for frog and cat muscles. Another motivation in obtaining a model for the dynamics of human muscle in situ is its more direct applications to two potentially rewarding areas, (1) the bioelectric control of prosthesis, where actual muscle signals are used to control movement and (2) the artificial stimulation of paralyzed limbs for patients with an upper motor lesion but whose muscles are in a normal state.

---

† For a more detailed presentation refer to Stanford Electronics Laboratories TR No. 6303-1, Stanford University, February 1970. This investigation was conducted under NASA Grant NGR-05-020-007 and Veterans Research Administration Grant RD-2152-M.

†† The author is now associated with AC Electronics Division of General Motors Corporation, Defense Research Laboratories.

The approach involved the formulation of isotonic and isometric experiments to obtain the explicit characterization of the mathematical equations for the postulated mechanism. The experimental procedure consisted of the artificial stimulation of the muscles under maximal and tetanic conditions via a surface electrode, application of external forces, and recording of the resulting motions. No voluntary activation of the muscles was tolerated. The external forces were applied to the forearm by means of an electric torque motor that was controlled not only as a torque-producing device but also as a regulator to compensate for developing inertial loads during accelerated motions as well as for the changing geometry of the arm/muscle system. The method of identification used for determining the unknown functions and parameters entering the equations of motion was purely empirical. Curve fitting the pertinent data recorded from the isotonic experiments permitted the explicit characterization of the contractile element. The results of the isometric experiments revealed the steady-state relationship between muscle torque and forearm position and, combined with the results of the isotonic experiments, allowed the characterization of the series elastic element.

CONCEPTUAL MODEL AND EQUATIONS OF MOTION

A. Conceptual Model Postulated

A mechanical analog-to-muscle force and kinematic behavior is the model postulated here. It is illustrated by Fig. 1 and consists of two elements in series: (1) a force generator, whose output depends on muscle excitation and length, in parallel with a damper and referred to as the contractile element and (2) an undamped elastic element called the series elastic element. The contractile and series elastic elements possibly could be mechanical equivalent elements resulting from actual parallel and series combinations of similar elements in the myofibrils. The elastic element could correspond to the thin filaments and the contractile element to the overlap of both thin and thick filaments. In the schematic representing the contractile element, F denotes the output of the force generator and is represented as a function of both muscle excitation $a(t)$ and of the change in muscle length x_L. It should be noted that the elements making up the contractile and series elastic elements are not implied to be linear; in fact, as this study will demonstrate, they are all nonlinear elements.

B. The Equations in General Form

1. Translational Equations and Coordinates for Skeletal Muscle

If the following translational generalized coordinates are defined,

$x_C \equiv$ distance shortened by contractile element from its normal resting length†

† "Normal resting length" of the muscles considered here is defined as that length corresponding to the smallest angle of forearm flexion for completely passive agonist and antagonist and for no tension in the muscle (i.e., for normal resting length $a(t)$, $P = 0$). θ_0 is approximately 40°.

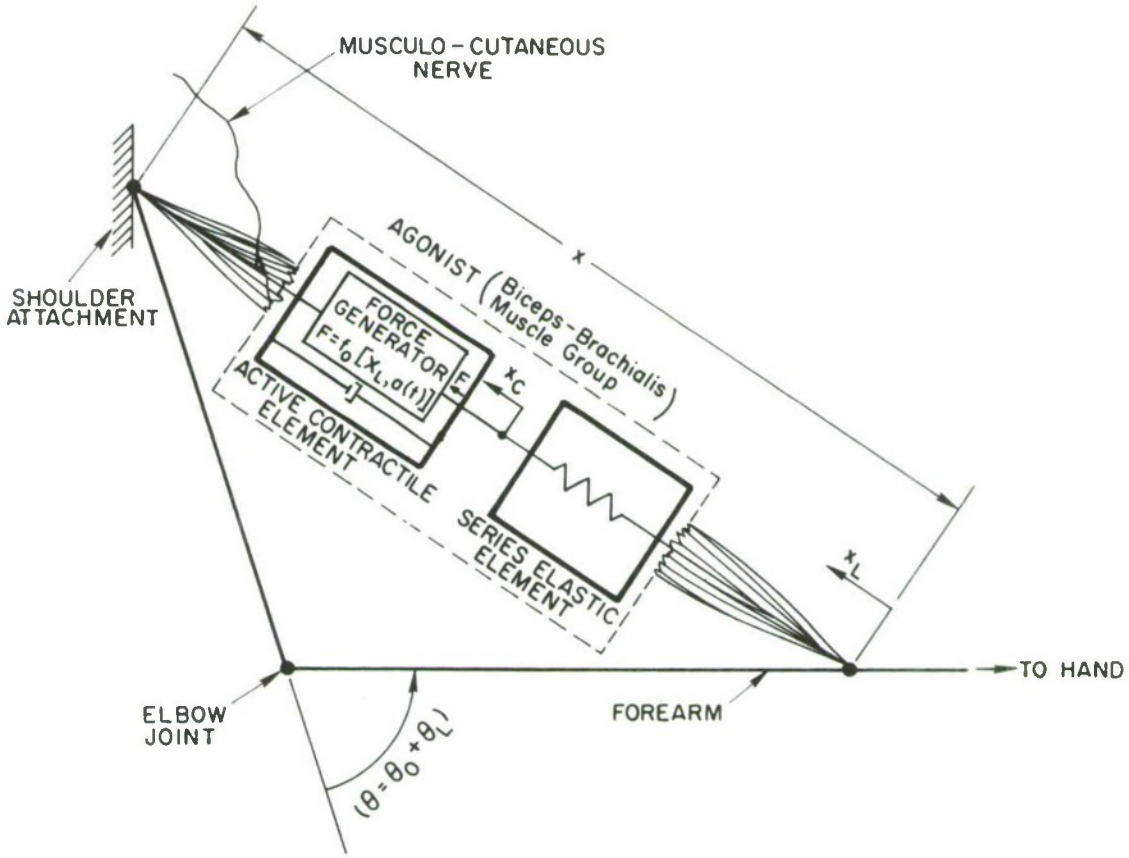


Fig. 1. MODEL OF TORQUE-PRODUCING MUSCLE FOR FOREARM FLEXURE. (Contractile and series elastic elements in series.)

x_L = distance contracted by muscle group from its normal resting length (i.e., distance moved by attachment point of muscle at forearm from normal resting-length position when shoulder attachment point is fixed)

P = tension in muscle

then, the equations characterizing the series elastic and contractile elements can be represented, respectively, in the following general forms:

$$x_C - x_L = f_1(P) \tag{1}$$

and

$$\dot{x}_C = f_2\left[P, x_L, a(t)\right] \tag{2}$$

2. Conversion of the Translational Equations into Rotational Forms

The recording of information is most convenient in rotational rather than translational coordinates and it is believed that the mathematical model

developed in this study will be more useful and of more direct application if expressed in rotational form. For these reasons, the transformation of the general translational equations into rotational forms is appropriate here.

If the following rotational parameters are defined as

θ_0 = angular position of arm, measured from fully stretched position, for normal resting length of muscle group

θ_C = angular displacement of forearm from its normal resting-length position θ_0 corresponding to contraction x_C if series elastic element is assumed fixed in length; i.e., if $x_C - x_L = 0$

θ_L = actual angular displacement of forearm from its normal resting-length position θ_0; i.e., angular displacement corresponding to x_L

then their relationships to the translational coordinates x_C and x_L, as derived in the Appendix are

$$x_i = \left(\ell^2 + d^2 + 2\ell d \cos \theta_0\right)^{1/2} - \left[\ell^2 + d^2 + 2\ell d \cos (\theta_0 + \theta_i)\right]^{1/2} \qquad (3)$$

where $i = C, L$. By using this equation, Eq. (1) can be rewritten as

$$\left[\ell^2 + d^2 + 2\ell d \cos (\theta_0 + \theta_L)\right]^{1/2} - \left[\ell^2 + d^2 + 2\ell d \cos (\theta_0 + \theta_C)\right]^{1/2} = f_1(P) \qquad (4)$$

The dynamic characteristics of the series elastic element, described by Eqs. (1) and (4), will not be characterized directly from an experimental procedure; instead, they will be characterized by making use of the mathematical model for the contractile element that will be developed directly from an experimental investigation. It will be apparent from the results, however, that the mathematical representation of θ_C is quite complex and, in turn, would render Eq. (4), which is the characterization of the series elastic element, very cumbersome and question its usefulness. For these reasons, it is desirable to simplify Eq. (4) to the following approximate form,

$$\theta_C - \theta_L = g_1(P) \qquad (5)$$

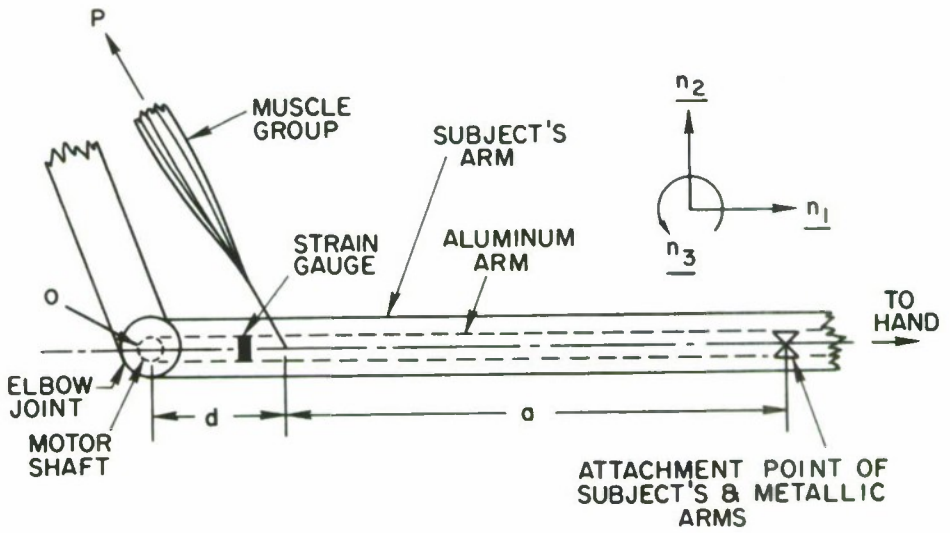
From the relationships of the Appendix, in particular Eqs. (A.9), (A.10), and (A.12), and for constant stimulation (maximal and tetanic) it is clear that the mathematical model for the contractile element, represented by Eq. (2), can be expressed as a general function of muscle-group torque about the elbow and of the rotational coordinates as

$$\dot{\theta}_C = g_2(T_A, \theta_L) \qquad (6)$$

3. Equations of Motion for Arm/Muscle/Electric-Motor System

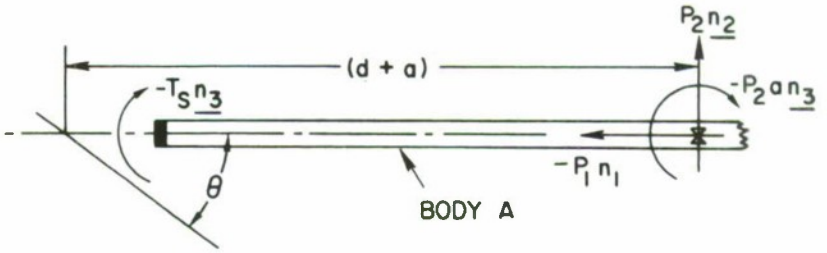
In the experimental method provision is made for the application of an external torque to the forearm by means of an electric motor. A brief description of the system is presented here to identify the interacting forces

entering the equation of motion. The forearm is strapped on an aluminum arm attached to the shaft of a torque motor, with the elbow coinciding with the center of the shaft. The subject is seated on a chair whose height can be so adjusted (dental chair) that flexion of the forearm takes place in a horizontal plane. The shoulders are securely strapped to the chair so that attachment points of the muscles at the shoulders can be considered fixed throughout the experiments. A measure of torque is provided by a strain gauge mounted on the aluminum arm. Free-body diagrams of the arrangement illustrating the pertinent forces acting on the system are shown in Fig. 2. These diagrams are not scaled



a. Schematic representation

b. Free-body diagram of metallic arm



c. Free-body diagram of body A, (metallic arm excluding portion between strain gauge and shaft)

Fig. 2. FORCES ACTING ON ARM/MUSCLE/ELECTRIC-MOTOR SYSTEM.

realistically, and the dimensions are exaggerated where appropriate for illustrative purposes.

Now, if

$\underline{n_i}$ = set of orthogonal unit vectors, where $i = 1,2,3$

O = rotation point or center of shaft and elbow joint

I_T = moment of inertia of all moving parts (metallic arm + forearm + motor-rotating parts) about O and along unit vector $\underline{n_3}$

I_A = moment of inertia of forearm + body A about O and along $\underline{n_3}$

$\underline{T_M}$ = electric-motor torque about shaft O

$\underline{T_S}$ = torque (bending moment) at strain-gauge location

$\underline{T_A}$ = muscle-group torque about elbow O

P_i = muscle-group force component along n_i vector

d = distance between muscle-group attachment point at forearm and elbow

a = distance between muscle-attachment and metallic-arm-attachment points at forearm

$\underline{H}$ = angular momentum about the fixed point O

$\Sigma\underline{M}$ = summation of the moments of all external forces about O

then the equation of rotational motion

$$\sum \underline{M} = \frac{d}{dt} (\underline{H})$$

for the arm/muscle/electric-motor system is (Fig. 2b)

$$-T_M\underline{n_3} - P_2 a\underline{n_3} + \left[(d + a)\underline{n_1} x \underline{P}\right] = \frac{d}{dt} (I_T \dot{\theta} \underline{n_3})$$

or

$$-T_M\underline{n_3} - P_2 a\underline{n_3} + \left[(d + a)\underline{n_1} x (P_1\underline{n_1} + P_2\underline{n_2})\right] = I_T \ddot{\theta} \underline{n_3}$$

which becomes

$$-T_M - P_2 a + (d + a)P_2 = I_T \ddot{\theta}$$

and

$$P_2 d = T_M + I_T \ddot{\theta} \qquad (7)$$

Noting that $\ddot{\theta} = \ddot{\theta}_L$ because $\theta = \theta_0 + \theta_L$ and θ_0 is a constant, and also that $P_2 d$ is the total torque of the muscle group about the elbow T_A because flexion of the arm takes place in a horizontal plane, Eq. (7) becomes

$$T_A = T_M + I_T \ddot{\theta}_L \tag{8}$$

The relationship between muscle-group torque T_A and muscle force P, given by Eq. (A.7), is rewritten here for completeness:

$$T_A = \frac{P \ell d \sin (\theta_0 + \theta_L)}{\left[\ell^2 + d^2 + 2\ell d \cos (\theta_0 + \theta_L)\right]^{1/2}} \tag{9}$$

Substituting (9) into (8), the following equation for muscle force is obtained:

$$P = (T_M + I_T \ddot{\theta}_L) h_5 (\theta_L) \tag{10}$$

where

$$h_5(\theta_L) = \frac{\left[\ell^2 + d^2 + 2\ell d \cos (\theta_0 + \theta_L)\right]^{1/2}}{\ell d \sin (\theta_0 + \theta_L)} \tag{11}$$

FORMULATION OF EXPERIMENTS

In Eqs. (5) and (6) the functions $g_1(\cdot)$ and $g_2(\cdot)$ are only symbolic. It is the purpose of the experimental procedure, based on the hypothesis of these equations, to determine the simplest forms of g_1 and g_2 to explain the data on muscle dynamic behavior.

It is important to realize that, of the generalized coordinates, θ_L (which is the angular displacement of the forearm from its normal resting-length position) is the only one that with its derivatives can be observed and measured experimentally. The coordinate θ_C, which corresponds to the shortening of the contractile element, cannot be observed directly. An important feature of the experimental procedure, therefore, will be the collection of data to permit the calculation of θ_C and $\dot{\theta}_C$ and the determination of the functions $g_1(\cdot)$ and $g_2(\cdot)$.

A. Characterization of the Contractile Element

The aim in describing the viscous or contractile element of the postulated model is to characterize $g_2(\cdot)$ or Eq. (6). The classic approach is through isotonic experiments, that is, experiments under which the muscle force P is kept constant throughout the duration of the experiment. Such experiments allow the determination of θ_C by observing θ_L. This can be established by taking the time derivative of Eq. (5)

$$\dot{\theta}_C - \dot{\theta}_L = \frac{\partial g_1(P)}{\partial P} \dot{P} \tag{12}$$

and by observing that, when P is constant, $\dot{P}$ is zero and $\dot{\theta}_C = \dot{\theta}_L$, or θ_C only differs from θ_L by a constant. This constant is the angular displacement of the forearm that would correspond to the stretch in the series elastic element from its normal resting length for the particular constant load P_C.

Without referring to the mathematical equations, the above observations can be made by considering only the mechanism of contraction of the postulated model (Fig. 1). If the muscle group is stimulated and if a constant load P_C is applied by the forearm† to the stimulated muscle, the contractile element will shorten, stretching the series elastic element in the process until the latter assumes a tension equal to the load P_C. This phase of the contracting process is isometric; that is, the total muscle length has not changed because the forearm is still in its initial position. What follows, however, is the isotonic process which is of concern here. Once the tension of the series elastic element has assumed the constant applied load P_C, any further shortening of the contractile element results in no change in the length of the series elastic element but rather in flexion of the forearm; both ends of the series elastic element move in such a way that its total length remains unchanged because its length, corresponding to a unique load P_C, is also unique. As a result, according to the definition of θ_C, the change in θ_C as measured from the beginning of the isotonic process (i.e., from the instant the forearm starts moving) equals θ_L‡, and the shortening velocity of the contractile element $\dot{\theta}_C$ equals $\dot{\theta}_L$ which is the actual observed angular velocity of forearm flexion. That is, under isotonic conditions,

$$\theta_C = \theta_L + \theta_{CK} \tag{13}$$

and

$$\dot{\theta}_C = \dot{\theta}_L \tag{14}$$

where θ_{CK} is the constant angular displacement of the forear, corresponding to the stretch of the series elastic element from its normal resting length for constant load P_C.

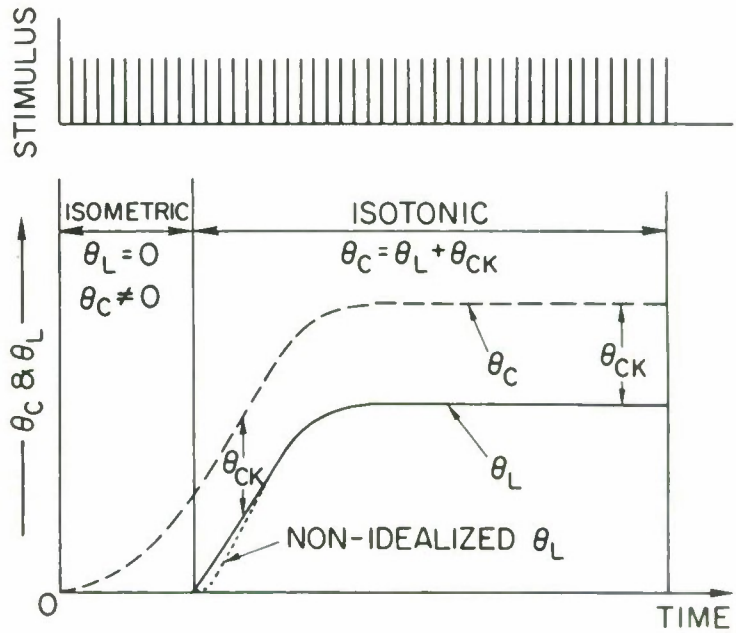
Observation of θ_L and $\dot{\theta}_L$ are, therefore, an indirect but accurate observation of θ_C and $\dot{\theta}_C$ during isotonic experiments and, because θ_C and $\dot{\theta}_C$ are not available for direct measurement, these experiments afford the best data for characterizing Eq. (6).

B. Description of Expected Results Based on the Postulated Model

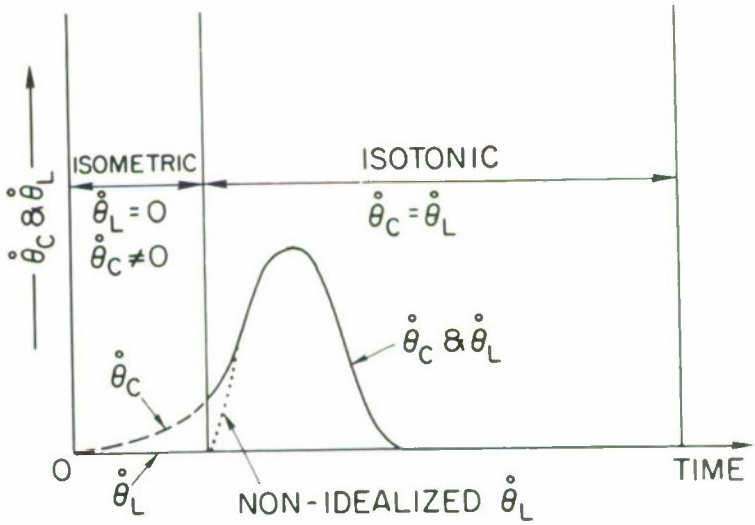
Figures 3 and 4 illustrate both the isometric and isotonic processes taking place in the experiments that characterize the contractile element. Motion of the forearm $(\theta_L, \dot{\theta}_L)$ only begins when the series elastic element has been stretched to the length corresponding to the tension P_C; that is, when the developed tension of the series elastic element equals the applied load. During

† The load is actually applied by a dc electric-torque motor with the forearm strapped to a metallic member that is keyed to the motor shaft. The load tends to move the forearm toward a stretched position.

‡ Similarly, the change in x_C equals x_L during the isotonic process.



a. Forearm angular position



b. Forearm angular velocity

Fig. 3. QUALITATIVE TIME HISTORIES OF θ_C, θ_L, $\dot{\theta}_C$, AND $\dot{\theta}_L$ FOR POSTULATED MODEL.

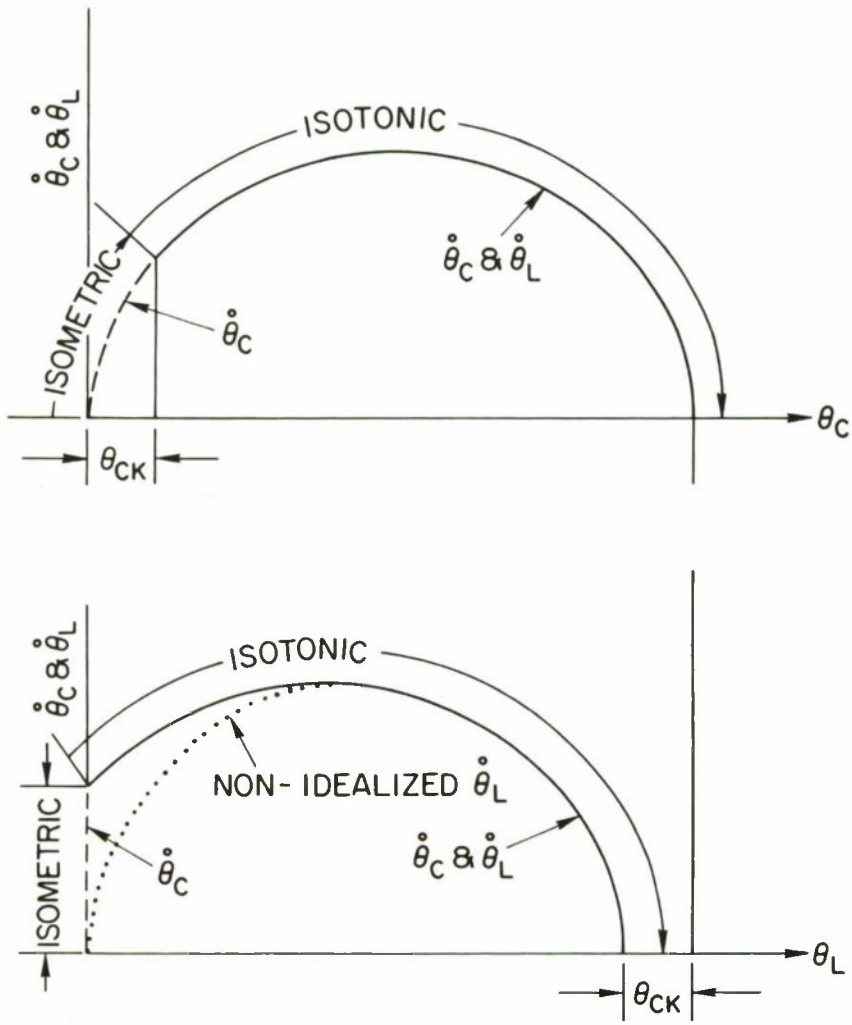


Fig. 4. QUALITATIVE PHASE-PLANE REPRESENTATIONS OF θ_C VS $\dot{\theta}_C$, $\dot{\theta}_L$ AND θ_L VS $\dot{\theta}_C$, $\dot{\theta}_L$.

the isotonic process, $\dot{\theta}_C = \dot{\theta}_L$ and θ_C always differs from θ_L by a constant θ_{CK} if the muscle is under constant tension. The magnitude of θ_{CK} depends on the value of the load applied to the forearm. Theoretically, the model suggests an instantaneous finite nonzero value for $\dot{\theta}_L$ at the onset of the isotonic process because $\dot{\theta}_L$ then is equal to $\dot{\theta}_C$, as illustrated in Figs. 3b and 4; however, this also implies an instantaneous velocity of the forearm at the beginning of motion. This cannot occur physically, and dotted lines are drawn to indicate the more realistic shapes of the expected curves under actual physical conditions.

It should be noted that Figs. 3 and 4 are not scaled realistically. For example, the isometric period, which only lasts approximately 75 ms, is greatly exaggerated relative to the total stimulus duration of 1300 ms. Also, the phase-plane plots of Fig. 4 are not quantitatively compatible to the time histories of Fig. 3; they are purely a working description of the postulated model.

They will prove very useful, however, in interpreting the experimental data required to characterize the function $g_2(\cdot)$.

C. Servo-Loop Implementation for Isotonic Experiments

1. The Problem

By considering Eq. (10), it is apparent that, if an experiment is to be isotonic (if P is to remain constant throughout), the right-hand side of the equation must be kept constant. The following two significant difficulties, however, complicate such isotonic experiments.

(a) The angular acceleration of the forearm deflection $\ddot{\theta}_L$ can be assumed by no means to be constant throughout such experiments; from the observations of past investigators, it is not. Therefore, unless the product $I_T\ddot{\theta}_L$ can be neglected (unless the inertia of the system I_T is negligibly small), some adjustment to compensate for the variations in $I_T\ddot{\theta}_L$ must be made for isotonic conditions to exist.

(b) The function $h_5(\theta_L)$ is not constant because of the changing geometry of the arm/muscle system during forearm flexion. In essence, $h_5(\theta_L)$ accounts for the changing length of the moment-arm vector $\underline{q}$ in the equations of motion (see Fig. 21 in the Appendix).

The inertia term I_T is, therefore, the only constant on the right-hand side of Eq. (10) because no change of mass occurs during the experiments. The applied load (electric-motor torque) T_M is the only other term that can be controlled easily and kept constant (by applying a constant input current to the motor), if the hope was to keep every individual parameter on the right-hand side of the equation constant. However, T_M will not be kept constant, and it is this flexibility that will permit true isotonic experiments to be conducted.

2. The Solution

The method of solution adopted was to use the motor not only as a torque-producing device to apply the load T_M to the forearm but also as a regulator in adjusting the value of T_M to compensate for the variables $I_T\ddot{\theta}_L$ and $h_5(\theta_L)$ in such a way as to keep the right-hand side of Eq. (10) constant.

The sero-control logic of Fig. 5 illustrates this solution. The input voltage V_i, corresponding to the desired constant force P_i in the muscles for a given isotonic trial, is kept constant throughout the experiment. The middle loop, involving the sensing of the forearm angular acceleration $\ddot{\theta}_L$ and the feeding back of the negative product $-I_T\ddot{\theta}_L$ to the summing junction, provides the solution for difficulty (a) by precisely compensating for the variable product $I_T\ddot{\theta}_L$. The outer loop, resulting in the feeding back of the inverted function $1/h_5(\theta_L)$ to a multiplier as shown, accounts for difficulty (b) which is the changing geometry of the arm/muscle system during flexion of the forearm. It should be observed that, without the outer loop, the system indicated by Fig. 5 would provide experimental trials with constant muscle torque rather than constant muscle force P; the right-hand side of Eq. (8)

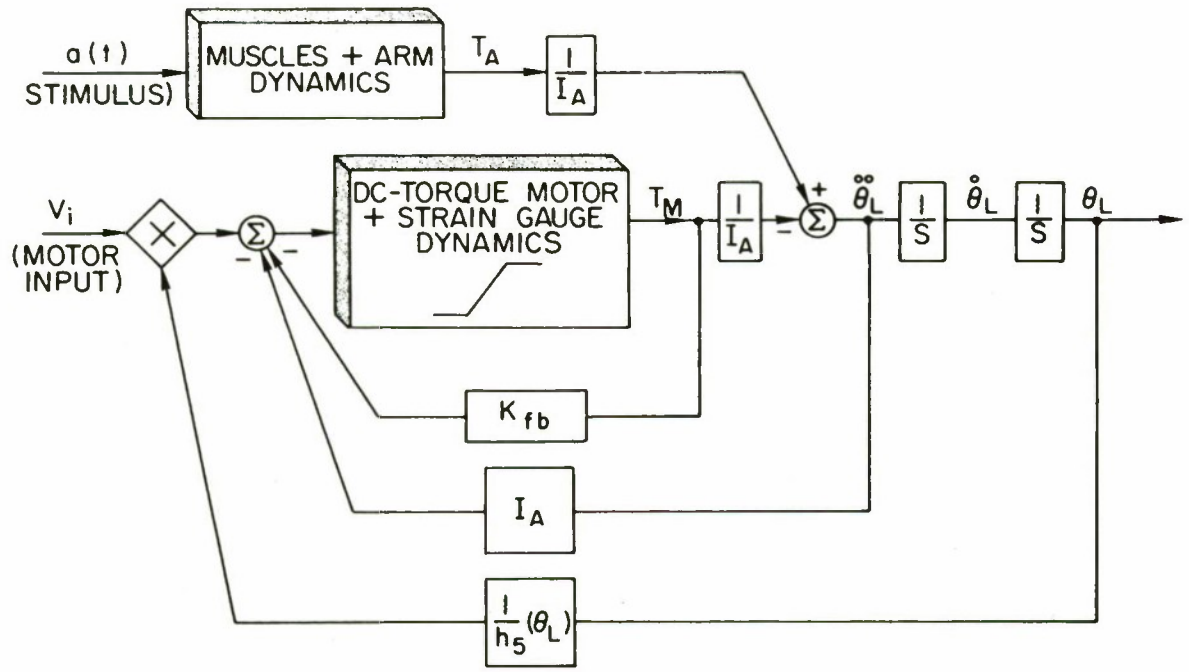


Fig. 5. BLOCK DIAGRAM OF ARM/MUSCLE/ELECTRIC-MOTOR SYSTEM FOR ISOTONIC EXPERIMENTS.

would have been kept constant during the trial rather than the right-hand side of Eq. (10). The inner loop (the feeding back of motor torque) is solely for the purpose of improving the dynamic response of the torque motor.

D. Characterization of the Series Elastic Element

Under isometric conditions, where the total muscle length remains fixed throughout the experiment or, for the case considered here, where both the shoulder and the forearm are restricted from moving,

$$T_A = T_S$$

Therefore, by fixing the metallic member on which the subject's arm is attached and by stimulating the muscle group, the time history of the torque T_A of the muscles about the elbow can be obtained by calculating T_S from the observed strain-gauge signal. Consequently, if the length of the muscle, or the position of the forearm, is varied for each trial, the forearm angular position as a function of the steady-state muscle torque can be established easily from the resulting data, namely

$$\theta_L = r(T_A) \tag{15}$$

Using these results, Eq. (6) can be expressed in terms of only two dependent variables; thus, by substituting Eq. (15) into (6) for θ_L,

$$\dot{\theta}_C = g_3(T_A) \tag{16}$$

Integration of (16) yields

$$\theta_C = \int_0^t g_3(T_A)\, dt \tag{17}$$

where t = real time.

Finally, by making use of Eqs. (15) and (17), the characterization of the series elastic element by Eq. (5) becomes, for steady state conditions,

$$\theta_C - \theta_L = \int_0^t g_3(T_A)\, dt - r(T_A) \tag{18}$$

subject to the constraint $\theta_L = r(T_A)$ of Eq. (15).

EXPERIMENTAL METHODS

A LINC-8 digital computer provided and controlled the input waveforms to the constant current stimulator (an ELS Model CCS-1A); it also monitored all the experiments and the recorded data by "marking" the analog tape on which the data are collected according to a "marker subroutine." Portions of the time histories recorded on tape were displayed instantaneously on an oscilloscope. All experiments were initiated by either the Stimulate & Mark (S/M) switch or the Stimulate (S) switch. The differences between these switches are that the S-switch only triggers the stimulator and does not initiate the "marker logic," and the S/M-switch does both. This choice of either "marking" or not marking the tape and not incrementing the run number of the experiment was found desirable because of the great number of unsatisfactory trials at the beginning of an experiment. A great many trials usually took place before locating the motor point† of the muscle group, determining the amplitude of the stimulus I required for maximal contractions, and obtaining consistency in the muscle torque. Because the data resulting from these preliminary trials were of no interest, no time or effort was devoted to digitizing these data; this was accomplished by using the S-switch, instead of the S/M-switch and, therefore, not marking the tape.

A block diagram of the overall system is provided by Fig. 6. It applies to both the isotonic and isometric experiments, noting that, in the isometric cases, the arm is fixed and the electric motor is not operative.

The subject's arm was attached to a metallic member that was keyed to the motor shaft, thus acting as a torque transducer. The method of connection between arm and transducer, as shown in Fig. 7, involved a small wrist cast, made for each individual subject, that was tightly secured to the transducer by means of a metal clamp. This allowed no play between arm and transducer. Elbow stops located and restrained the elbow-joint axis in a position coinciding with the motor-shaft axis.

---

† "Motor point" is defined as that electrode location on the surface of the skin that, on stimulation, produces the greatest isometric torque to develop about the elbow.

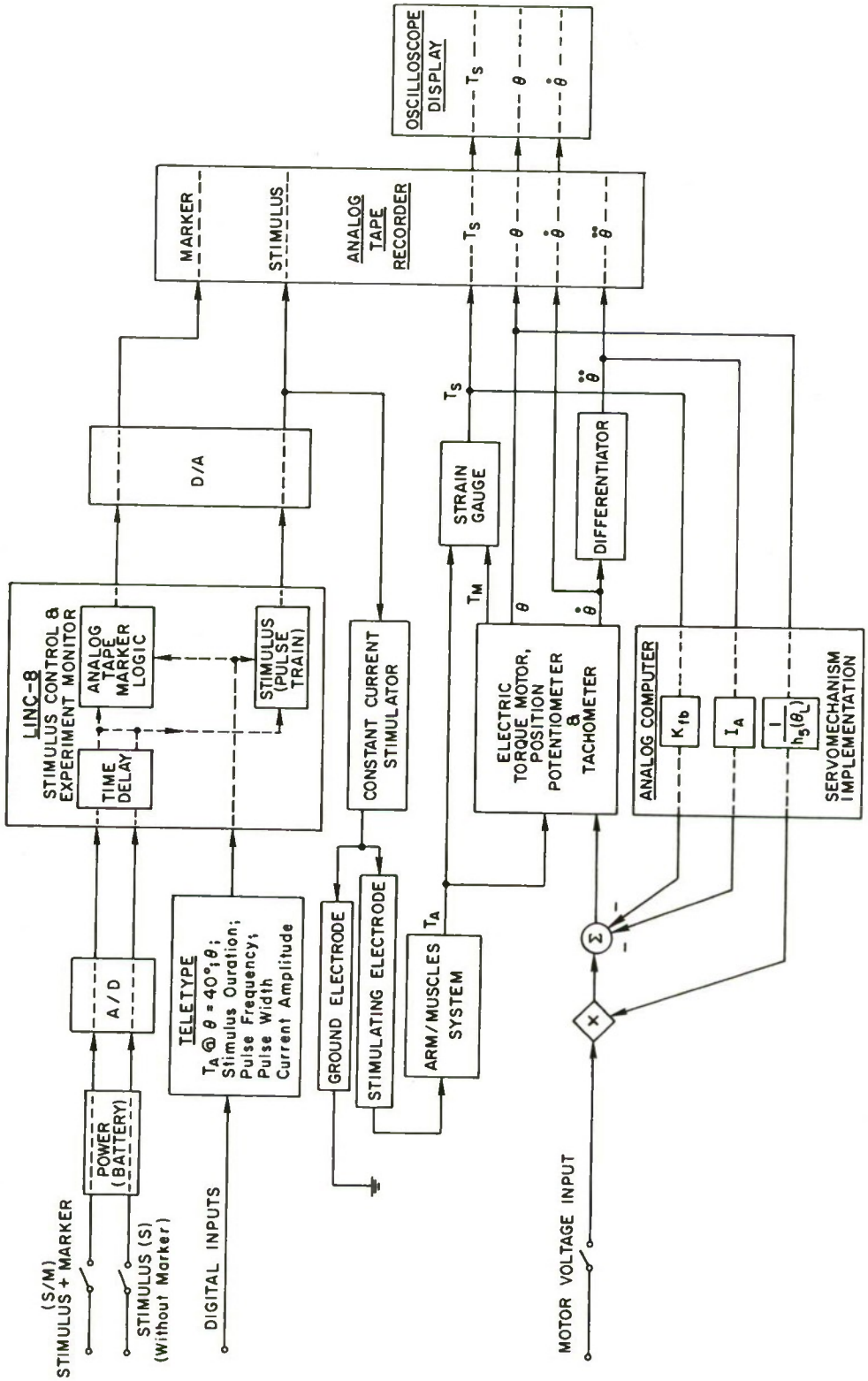


Fig. 6. BLOCK DIAGRAM OF EXPERIMENTAL SYSTEM.

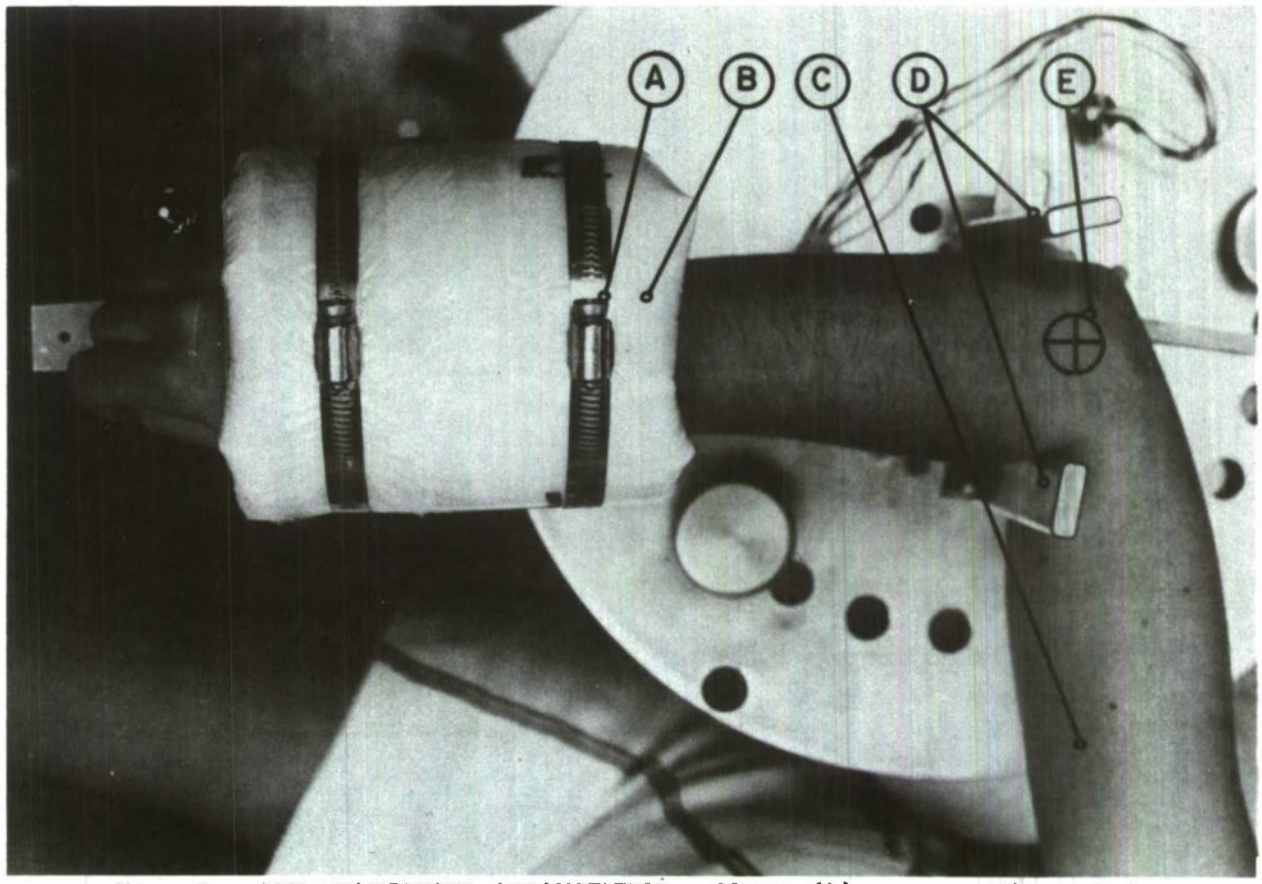


Fig. 7. ARM-TRANSDUCER ARRANGEMENT. Clamp (A) secures the wrist cast (B) to the torque transducer beam. The wrist cast, a polyurethane foam mold of the distal forearm and wrist joint, was prepared for each subject. The mold accurately reflected the contour of the limb and immobilized the wrist joint. The cast can be rotated within clamp (A) to adjust the radio-ulnar joint position. The biceps muscle lies in the frontal aspect of the upper arm (C). Elbow stops (D) locate and restrain the elbow-joint axis (E), holding it coincident with the torque transducer axis.

In the isometric experiments, the transducer or metallic arm is kept fixed by a metal pin through the transducer and the motor platform. In the isotonic experiments, the arm is free and the experiments are initiated with the subject's arm completely relaxed and the motor producing a torque tending to stretch the arm. This torque pushes the subject's arm and the transducer to rest against a stop (plastic peg) that is so positioned that the flexor muscles of the forearm are at their normal resting length (i.e., $\theta = \theta_0 \doteq 40°$).

Attitude and cooperation were an important part of the experiments because no activation of the muscle group from the central nervous system, voluntary or involuntary, could be tolerated. As a check for possible interference by the central nervous system, or for a slipping or misplaces timulating electrode, and for any other factor affecting the experimental results, the following precautions were taken.

(1) All trials were repeated until the time histories for strain-gauge torque T_S, in the isometric experiments, and forearm position θ, in the isotonic experiments, would be repeatable (overlapping curves) for at least three consecutive trials.

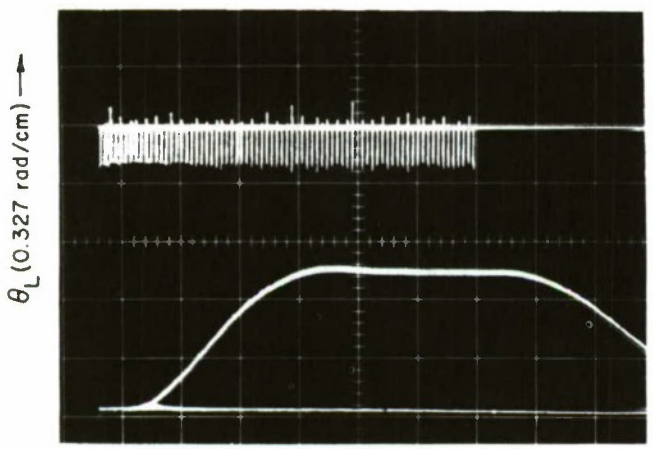
(2) After all the required data were collected, some of the initial runs were repeated as a "spot check" and compared for consistency. Polaroid photographs of the time histories selected for the spot checks provided immediate comparison.

EXPERIMENTAL RESULTS

The figures presented in this section are reproductions of actual photographs of the raw data, as they appear on an oscilloscope.

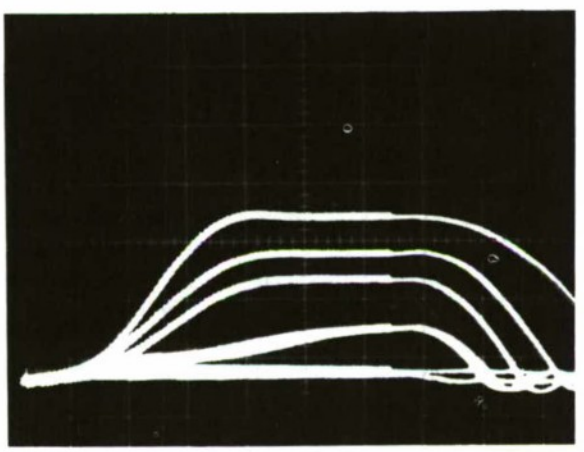
A. Isotonic Motions

Typical time histories of forearm flexion under isotonic conditions are provided by Figs. 8 and 9; the external loads applied to the forearm in the



Stimulus
Duration: 1280 ms
Amplitude: 8.0 mA
Frequency: 60/sec
Pulsewidth: 0.5 ms

Fig. 8. θ_L VS t; T_M = 1 FT-LB AT $\theta = \theta_0$.



Duration: 1280 ms
Amplitude: 8.0 mA
Frequency: 60/sec
Pulsewidth: 0.5 ms

Fig. 9. θ_L VS t; T_M = 1,2,3, 5,7 FT-LB AT $\theta = \theta_0$.

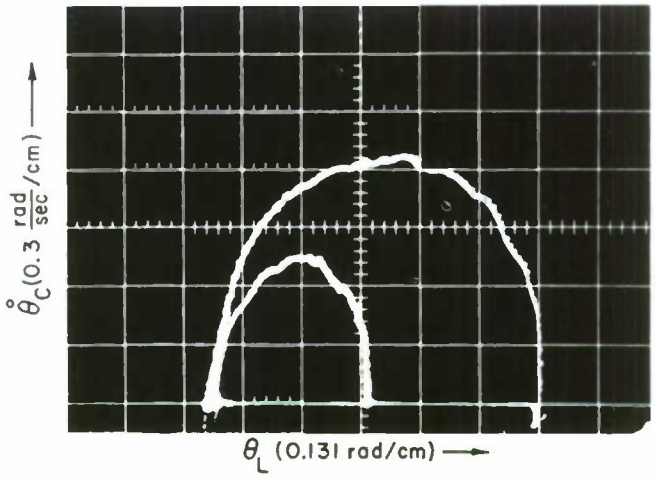
various trials are indicated on the figures. It is important to remember, however, that these external loads, applied by the electric motor are functions of both the angular position θ and angular acceleration $\ddot{\theta}$ of the forearm† and that the loads given in the figures apply to the initial conditions of each trial, namely when $\theta = \theta_0$ and $\ddot{\theta} = 0$. Figure 8 also includes the stimulus

† Refer to the control logic of Fig. 5 where continuous corrections to the load T_M are made according to the forearm angular position and acceleration.

train of pulses (60/sec) which allows the study of muscle contraction relative to its activation. All traces are triggered by the stimulus, and it is clear from both figures that there is a time delay between the onset of stimulation and the beginning of arm motion. This is not surpirsing because it was predicted by and accounted for by the mechanism of the postulated model; in fact, this time delay that varies between 100 and 200 ms, depending on the applied load, corresponds to the isometric period illustrated by Figs. 3 and 4. Some interference of the stimulus with the traces of Fig. 9 can be observed. These traces appear thicker when the stimulus is present and have the helpful, although unplanned, effect of illustrating its duration.

Both Figs. 8 and 9 also indicate a time delay between the end of activation and the beginning of relaxation of the muscles. This delay is about 60 ms. As soon as relaxation begins, the muscle torque can no longer match the applied external load which then brings the forearm back toward its original position until the transducer comes to rest against the mechanical stop. Some bouncing of the arm/transducer system against the stop occurs, as can be observed at the end of the traces of Fig. 9. Indeed, the only data of interest in this investigation are those collected prior to relaxation. Figure 9 shows traces for various loads and indicates that the isometric tension was not high enough to move the largest applied load of 7 ft-lb.

As discussed earlier, the elements comprising the postulated model are expected to be nonlinear which suggests analysis of the data in the phase plane. Typical phase-plane displays of the data for two different loads are illustrated in Fig. 10. By the definition of "phase plane," these are plots of the coordinate vs its first derivative, namely θ_L vs $\dot{\theta}_L$. Under isotonic conditions $\dot{\theta}_C = \dot{\theta}_L$, and the ordinate can also be labeled $\dot{\theta}_C$ because the purpose of the isotonic experiments is precisely to relate $\dot{\theta}_C$, θ_L, and T_A mathematically



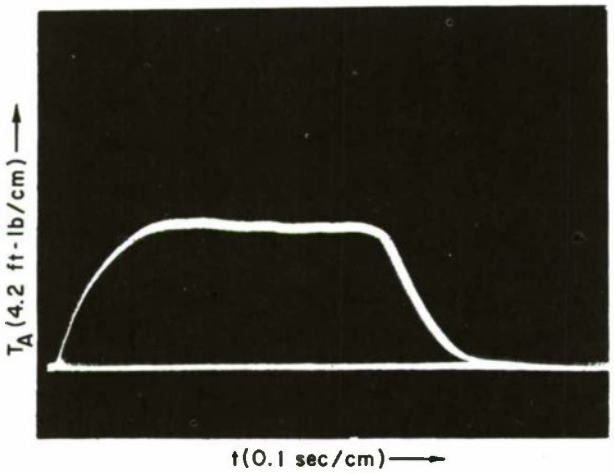
Duration: 1280 ms
Amplitude: 8.0 mA
Frequency: 60/sec
Pulsewidth: 0.5 ms

Fig. 10. $\dot{\theta}_C$ VS θ_L; T_M = 3,5 FT-LB AT $\theta = \theta_0$.

to characterize the contractile element. It will be shown later how phase-plane displays of the data, similar to those of Fig. 10, will form the basis for characterizing the contractile element.

B. Isometric Torque

The information recorded from the set of isometric experiments is typically illustrated by Fig. 11. The data consist of time histories of the developed



Duration: 500 ms
Amplitude: 8.0 mA
Frequency: 60/sec
Pulsewidth: 0.5 ms

Fig. 11. T_A VS t; $\theta_L = 30°$ (FOUR REPETITIONS).

isometric torque T_A for various angular positions of the forearm θ_L, that is, for various muscle lengths. Because no motion of the arm is involved in these experiments, they are considerably easier to conduct and the results obtained can be amazingly consistent and repeatable. Figure 11 is actually a photograph of four traces resulting from four consecutive trials that were recorded on an oscilloscope in the storage mode, and the traces were triggered by the stimulus. An almost perfect overlap of these time histories can be observed.

The stimulus duration for the trials was 500 ms and, again, it can be observed that the stimulus leads the muscle torque by a short time delay (20 ms); indeed, this delay need not equal the one preceding the start of motion in the isotonic experiments. In the isotonic case, the length of the delay is the time required for the developed muscle torque to equal the applied load and, therefore, depends on the value of the applied load; in the isometric data, the delay preceding the start of muscle torque buildup is due solely to physiological factors. The delay observed between the end of stimulation and muscle relaxation, however, should be the same in both the isotonic and isometric trials because it depends only on physiological factors. This appears to be the case because, again, the end of the period of stimulation leads the beginning of muscle relaxation by about 60 ms.

CHARACTERIZATION OF EQUATIONS FROM EXPERIMENTAL DATA

The chronological steps taken in this investigation and presented in the preceding sections can be summarized as follows:

(1) postulation of a dynamical mechanism as a model for human skeletal muscle

(2) mathematical description of the model by equations in general forms involving undetermined functions and parameters

(3) formulation of experiments that would permit the identification of the unknown functions and parameters in the equations

(4) experimental procedures and collection of data for the explicit characterization of the undetermined functions and parameters

To obtain a complete mathematical model for the dynamics of the muscles studied, the remaining step is to make use of the data collected to characterize fully the unknown functions $g_1(\cdot)$ and $g_2(\cdot)$ of the model and to determine the parameters comprised by them.

A. The Functions and Parameters Characterizing the Contractile Element

The approach taken to explicitly describe Eq. (6) from the data collected is empirical; the purpose was to determine a general class of curves that would best fit the experimental data. The analog data was digitized on a LINC-8 computer and displayed on an oscilloscope, which provided great flexibility in manipulating the data in the search for curves that would best describe the results.

1. General Form of the Velocity-Length Function $g_2(\cdot)$

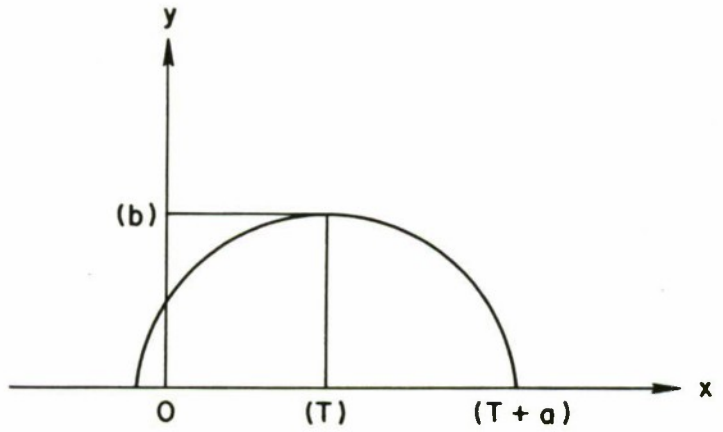
It was observed that the arches of the phase-plane plots were elliptical; distances from the center of ellipses to their major and minor vertices (a and b, respectively) could be selected as well as translations of the ellipses along their major axes so as to match the phase-plane plots. Indeed, only the portion of the ellipses that lies in the upper-right quadrant of the coordinate system is of concern.

This development can be better understood by beginning with the general equation for an ellipse, in terms of the familiar coordinates x and y:

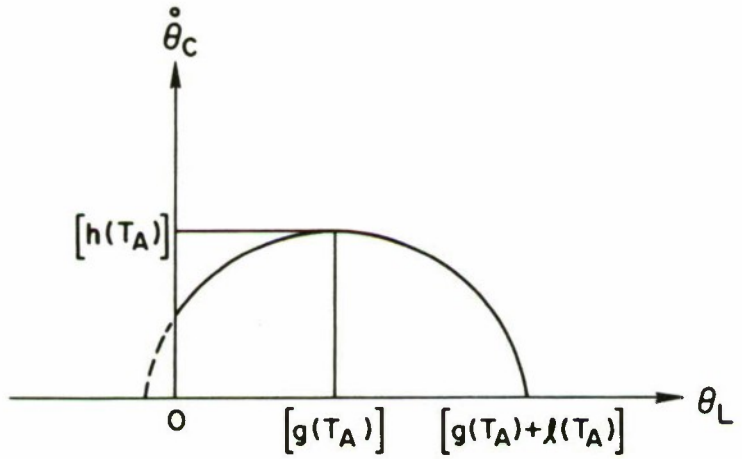
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \tag{19}$$

where a and b are its semimajor and semiminor axes lengths, respectively. If the center of the ellipse is translated a distance T along the abscissa as shown in Fig. 12a, the equation in the upper-half plane becomes

$$y = b\left[1 - \left(\frac{x - T}{a}\right)^2\right]^{1/2} \tag{20}$$



$$a. \quad y = b\left[1 - \left(\frac{x - T}{a}\right)^2\right]^{1/2}$$



$$b. \quad \dot{\theta}_C = h(T_A)\left\{1 - \left[\frac{\theta_L - g(T_A)}{\ell(T_A)}\right]^2\right\}^{1/2} \quad ; \quad \theta_L > 0, \quad \dot{\theta}_C > 0$$

Fig. 12. ELLIPTICAL DESCRIPTION OF ISOTONIC DATA.

Because a, b, and T are selected to have ellipses match the phase-planes corresponding to various muscle torques T_A, these parameters can be expressed as functions of T_A and Eq. (20) then can be written as

$$\dot{\theta}_C = h(T_A)\left\{1 - \left[\frac{\theta_L - g(T_A)}{\ell(T_A)}\right]^2\right\}^{1/2} \tag{21}$$

where x and y of Eq. (19) have been replaced by the coordinates θ_L and $\dot{\theta}_C$, and a, b, T, by the functions $\ell(T_A)$, $h(T_A)$, and $g(T_A)$, respectively. The association of these functions with the phase planes is illustrated by Fig. 12b. From Eq. (21), the velocity-length function, in turn, can be expressed as

$$g_2(T_A, \theta_L) = h(T_A)\left\{1 - \left[\frac{\theta_L - g(T_A)}{\ell(T_A)}\right]^2\right\}^{1/2} \tag{22}$$

The above method of describing the functions $\ell(T_A)$, $h(T_A)$, and $g(T_A)$ would be equivalent to considering a nominal elliptical arch and rescaling it about its major axis according to $\ell(T_A)$ and about its minor axis according to $h(T_A)$ after having shifted the ellipse by $g(T_A)$. This interpretation led to the description of $\ell(T_A)$ as the position-gain function, $h(T_A)$ as the velocity-gain function, and $g(T_A)$ as the translational function. These functions will be characterized explicitly in the following sections.

2. Characterization of the Position Gain Function $\ell(T_A)$

As discussed previously and indicated on Figs. 3 and 4, the initial portion of the data in all isotonic trials cannot accurately represent the dynamic behavior of the postulated model because the postulated mechanism implies a nonzero velocity of the forearm at the onset of isotonic motion, that is, when motion of the forearm is first observed. For this reason, only the latter parts of the phase-plane curves were used in extracting the data and, when necessary, they were extrapolated back to the initial regions of larger muscular lengths to appear as in Fig. 12b, that is, with $\dot{\theta}_C > 0$ at $\theta_L = 0$.

The position-gain function was derived from the data by tabulating the distances from the center of the extrapolated ellipses (phase planes) to their major vertices (those on the θ_L axis) and the corresponding flexor torque T_A. This information is given in Table 1 for three individual subjects. Curve fitting this data was attempted by displaying the points on linear semilog and log-log plots. Figure 13 illustrates the linear fit adopted and defined by the parameters of

$$\ell(T_A) = c - nT_A \tag{23}$$

The values for c and n are constants for each individual and are given in Table 2 for five subjects.

3. Characterization of the Velocity-Gain Function $h(T_A)$

The procedure here is similar to the one used to characterize the position-gain function. The distances from the center of the ellipses to their minor vertices and the corresponding flexor torque T_A are tabulated for the various trials. In this case, the best fit of the data was obtained by plotting $h(T_A)$ vs the adjusted torque T_A on semilog scales (Fig. 14). The resulting characteristic equation is

$$\log_{10} h(T_A) = -0.434\, aT_A + \log_{10} U \tag{24}$$

Table 1

THE TRANSLATIONAL, POSITION-GAIN, AND VELOCITY-GAIN FUNCTIONS

$$\dot{\theta}_C = g_2(\theta_L, T_A)$$

$$\dot{\theta}_C = h(T_A)\left\{1 - \left[\frac{\theta_L - g(T_A)}{\ell(T_A)}\right]^2\right\}^{1/2}$$

T_A at $\theta = \theta_0$ (ft-lb)	Subject 1			Subject 2			Subject 3		
	$g(T_A)$ (rad)	$\ell(T_A)$ (rad)	$h(T_A)$ (rad/sec)	$g(T_A)$ (rad)	$\ell(T_A)$ (rad)	$h(T_A)$ (rad/sec)	$g(T_A)$ (rad)	$\ell(T_A)$ (rad)	$h(T_A)$ (rad/sec)
0.1	0.524	0.576	2.22	0.366	0.418	1.8	---	---	---
2.0	0.401	0.420	1.50	0.288	0.302	1.2	0.372	0.486	1.86
3.0	0.345	0.399	1.31	0.196	0.236	0.96	0.300	0.472	1.59
4.0	0.305	0.367	1.08	---	---	---	---	---	---
5.0	0.236	0.275	0.94	0.106	0.107	0.64	0.214	0.428	1.11
6.0	0.197	0.199	0.78	---	---	---	0.184	0.310	0.84

or

$$h(T_A) = U\, e^{-aT_A} \tag{25}$$

Again, the tabulation of the data and of the resulting parameters are given in Tables 1 and 2, respectively.

4. Characterization of the Translational Function $g(T_A)$

The data of interest in this characterization are the translation of the phase planes along the abscissa for various loads; therefore, the data representing the distances between the center of the ellipses (phase planes) and the origin of the coordinate system were tabulated vs muscular torque in Table 1. A linear fit of these data is shown by Fig. 15 and described by

$$g(T_A) = b - mT_A \tag{26}$$

The values of b and m are given in Table 2 for five individuals.

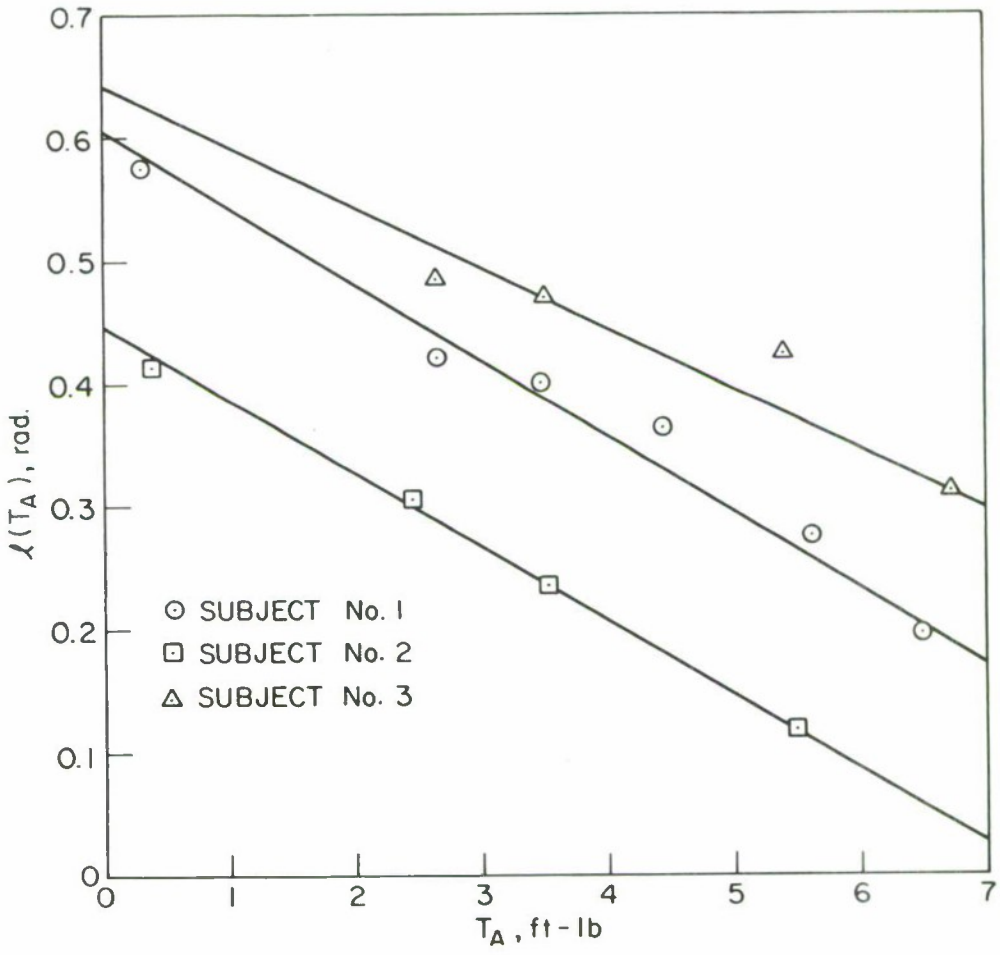


Fig. 13. POSITION-GAIN FUNCTION.

The results of this section now enable Eq. (6) to be explicitly described. In particular, substitution of Eqs. (23), (25), and (26) into Eq. (21) yields the complete mathematical model describing the dynamic behavior of the contractile element given as

$$\dot{\theta}_C = U\, e^{-aT_A} \left[1 - \left(\frac{\theta_L + mT_A - b}{c - nT_A} \right)^2 \right]^{1/2} \tag{27}$$

B. The Functions and Parameters Characterizing the Series Elastic Element

The purpose of this section is to explicitly describe Eq. (5). To this end, the results of both the isometric and isotonic experiments will be used in conjunction. Table 3 is a tabulation of the maximum torque developed T_A (steady state) for various forearm positions in the isometric experiments.

Table 2

PARAMETERS FOR THE CONTRACTILE-ELEMENT EQUATION

$$\dot{\theta}_C = g_2(\theta_L, T_A)$$

$$\dot{\theta}_C = U\, e^{-aT_A}\left\{1 - \left[\frac{\theta_L + mT_A - b}{c - nT_A}\right]^2\right\}^{1/2}$$

Equation Parameters	Subject				
	1	2	3	4	5
U (1/sec)	2.34	1.97	2.88	2.80	2.40
a	0.166	0.210	0.173	0.167	0.214
b	0.550	0.400	0.512	0.510	0.505
c	0.605	0.445	0.640	0.535	0.635
m (1/ft-lb)	0.0531	0.0543	0.0541	0.081	0.0457
h (1/ft-lb)	0.0615	0.0610	0.049	0.0776	0.0740

The approach taken to relate T_A and θ_L mathematically to develop the torque-length function is the same as that used to characterize the position-gain, velocity-gain, and translational functions. Figure 16 indicates a linear fit of the data, giving rise to the following torque-length function for steady-state conditions,

$$\theta_L = r(T_A) = d - kT_A \tag{28}$$

The parameters k and d assume constant values for each individual and are given in Table 4.

If θ_L is eliminated from Eq. (27), making use of Eq. (28), $g_3(T_A)$ becomes

$$g_3(T_A) = U\, e^{-aT_A}\left\{1 - \left[\frac{T_A(m - k) + d - b}{c - nT_A}\right]^2\right\}^{1/2} \tag{29}$$

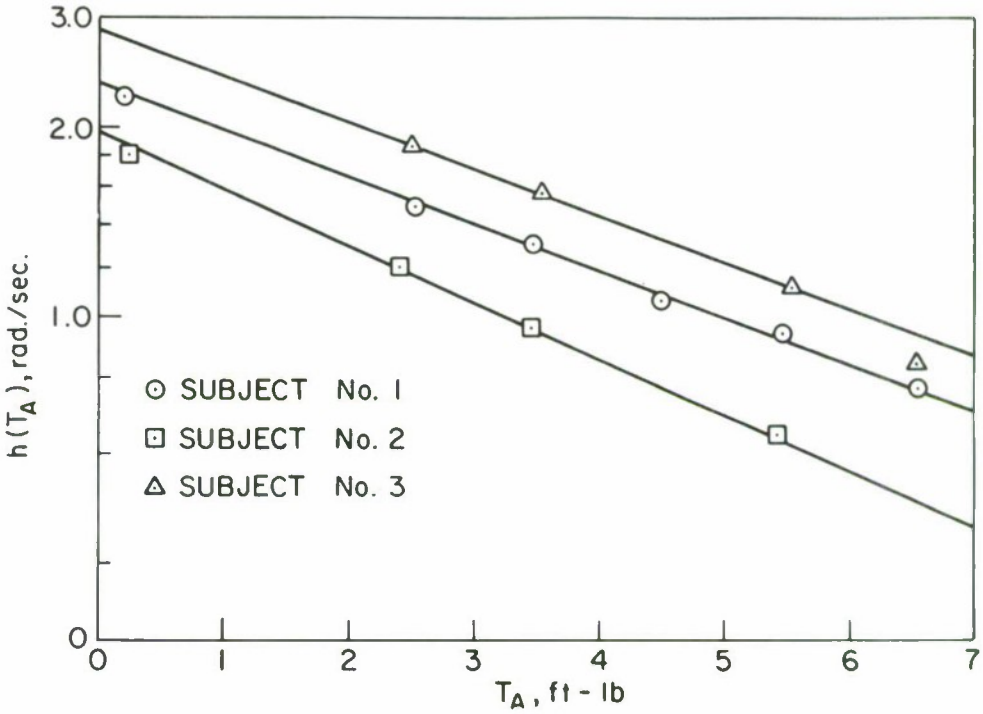


Fig. 14. VELOCITY-GAIN FUNCTION.

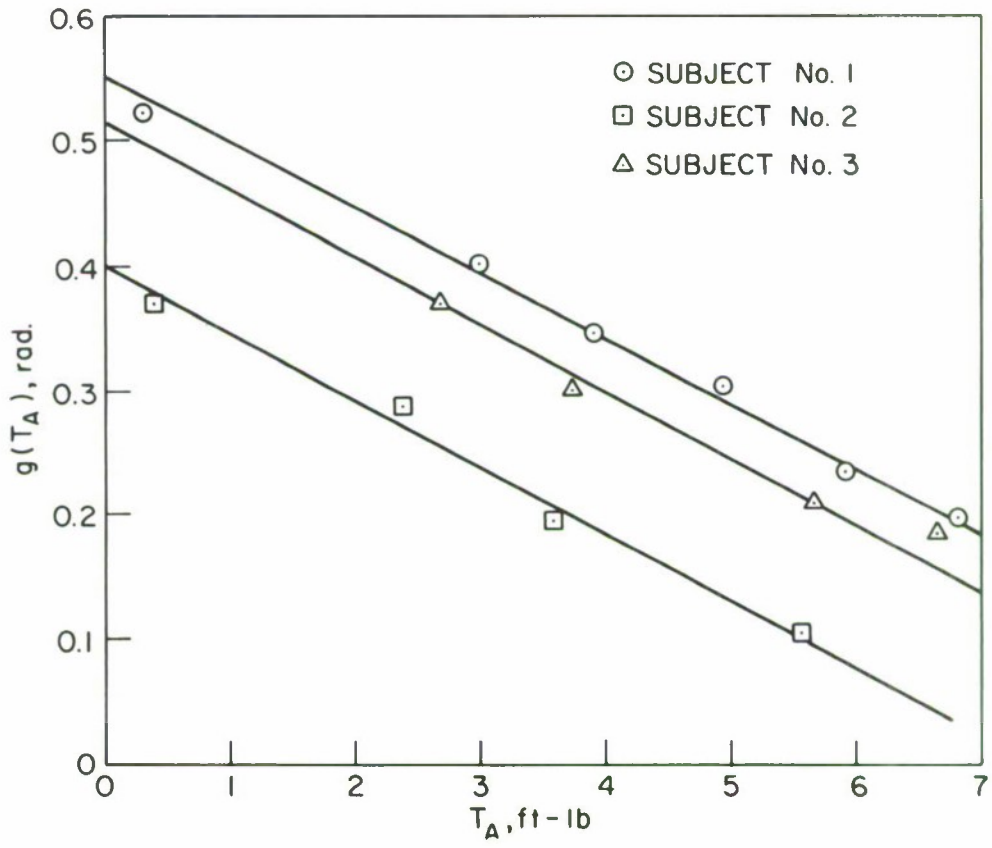


Fig. 15. TRANSLATIONAL FUNCTION.

Table 3

THE TORQUE-LENGTH FUNCTION

$$\theta_L = r(T_A)$$

$\theta_L + \theta_0$ (deg)	θ_L (deg)	θ_L (rad)	T_A (ft-lb)		
			Subject 1	Subject 2	Subject 3
45	5	.0873	---	12.60	13.84
60	20	.3491	14.08	10.48	11.83
75	35	.6109	11.52	8.40	10.44
90	50	.8727	10.08	6.72	9.36
105	65	1.1345	8.28	5.04	6.12
120	80	1.3963	6.12	2.52	5.04

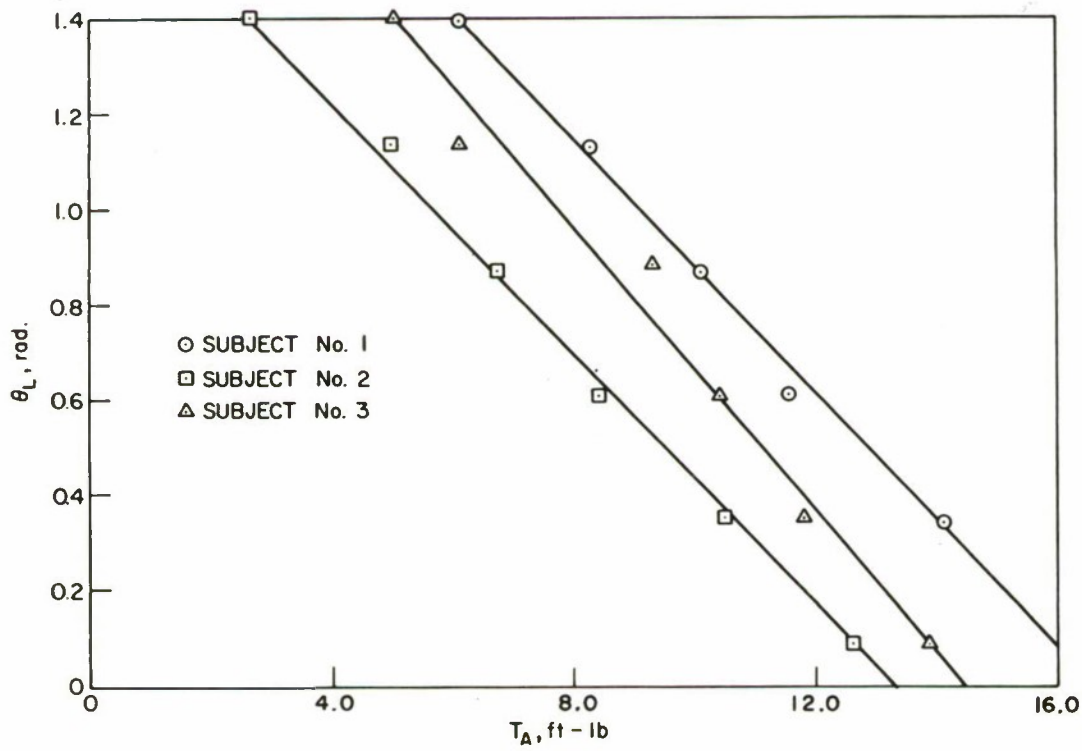


Fig. 16. TORQUE-LENGTH FUNCTION.

Table 4

PARAMETERS FOR THE TORQUE-LENGTH FUNCTION

$$\theta_L = d - kT_A$$

Equation Parameters	Subject		
	1	2	3
k (1/ft-lb)	0.134	0.130	0.147
d	2.21	1.75	2.13

and Eq. (18), which characterizes the series elastic element, is fully described.

C. Relationships between the Functions Characterizing the Contractile and Series Elastic Elements

From the definitions of the translational, position-gain, and torque-length functions, it is important to note the following relationship existing between them:

$$g(T_A) + \ell(T_A) = r(T_A) \tag{30}$$

This can be observed by referring to Fig. 12b, where the label $g(T_A) + \ell(T_A)$ indicates the forearm position at which equilibrium occurs. At this equilibrium position, the muscle is clearly isometric and the applied load (by electric motor) duplicates the resistance offered by the mechanical device (metal pin) restraining the motion of the metallic arm in the isometric experiments. As a result, the relationship between muscle length (forearm position) and muscle torque $\theta_L = r(T_A)$ in the isometric experiments and $\theta_L = g(T_A) + \ell(T_A)$ when equilibrium is reached in the isotonic experiments must be the same. By making use of Eqs. (23), (26), and (28), the relationship (30) can be rewritten as

$$(c - nT_A) + (b - mT_A) = d - kT_A$$

and it follows that

$$b + c = d \tag{31}$$

and

$$m + n = k \tag{32}$$

Relationships (31) and (32) provide useful checks for the experimental data and for the numerical values of the parameters entering the mathematical model.

The maximum muscle torque for a given forearm position can be obtained from the constraint equation (28) which relates steady-state torque to forearm flexion, as

$$T_{A_{max}} = \frac{d - \theta_L}{k} \tag{33}$$

and it now can be observed from Eqs. (31), (32), and (33) that an equivalent constraint equation for maximum muscle force also can be obtained from the results of the isotonic experiments, namely

$$T_{A_{max}} = \frac{(b + c) - \theta_L}{m + n} \tag{34}$$

These constraints show the dependency of maximum muscle force on muscle length and must be taken into consideration when studying the voluntary control of movement.

DISCUSSION

A. Evaluation of Model

The model developed from the results of the isotonic experiments, which constitutes the basis for the descriptions of both the contractile and series elastic elements, is evaluated in this section. Phase-plane plots computed from the derived model, Eq. (27), are compared with actual collected data, and the results are illustrated in Fig. 17 where the evaluation is made for three values of muscle force. Such a comparison is not as straightforward as it appears because, again, the change in geometry of the arm/muscle system complicates the evaluation. It must be remembered that the data comprising the phase-plane plots as displayed are for constant muscle force and not for constant muscle torque. Accordingly, adjustments in the value of muscle torque must be made for each corresponding forearm position when computing phase planes from the model.

As expected, the model predicts faster contractions of the contractile element at the onsets of forearm motions than the actual observed muscular contractions. This is compatible with the model postulated for the muscle as a whole and was discussed earlier. Namely, under real physical conditions, the velocity of the observed muscular contraction cannot instantaneously match the nonzero velocity of the contractile element at the onset of motion, as predicted by the model. A small time delay is required for these velocities to meet the theoretical prediction of their equality under isotonic conditions.

B. Comparison of Model

The best known mathematical description of muscle behavior was derived by Hill [Ref. 1] and given by

$$(P + \alpha)(V + \beta) = \beta(P_o + \alpha) \tag{35}$$

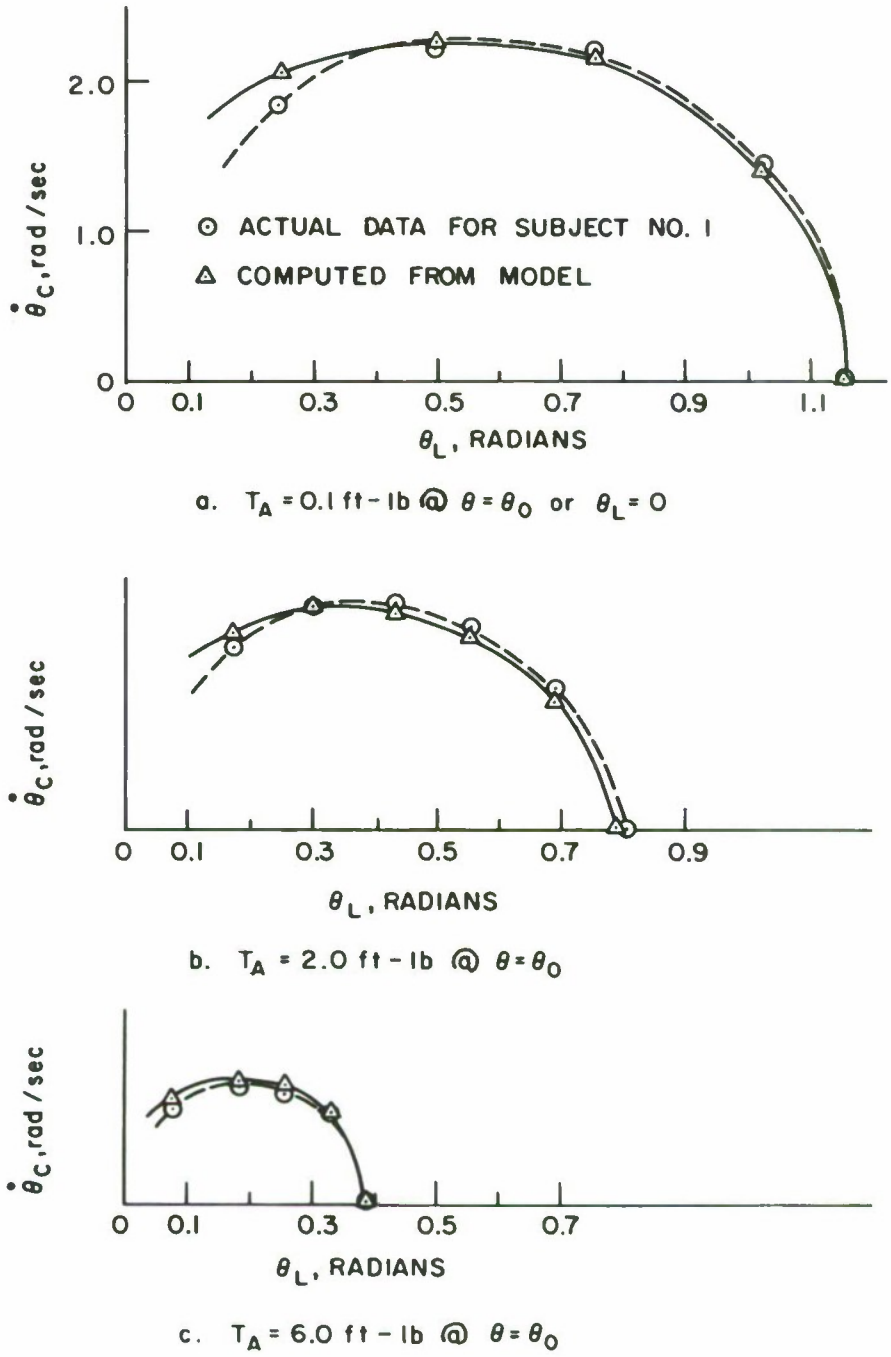


Fig. 17. COMPARISON OF CONSTANT-FORCE THEORETICAL AND EXPERIMENTAL CURVES.

where

P = external force

V = shortening velocity of the muscle

P_o = maximum isometric tension

α, β = constants

The principal limitation of this model is that it does not account for the length of the muscle, which was subsequently found to be a definite factor influencing the velocity and the tension in muscles [Refs. 8, 9, 10, and 4]. Hill's equation, therefore, relates muscle force and shortening velocity only at the studied muscular length which was the "normal resting length." This is the most notable distinction between Hill's model and the model developed in this investigation, Eq. (27), where the contractile-element velocity is explicitly described as a function of both muscle force and length.

Although Hill's experiments were conducted on the frog's sartorius muscle *in vitro* at 0°C, it is interesting to compare our results for human muscle *in situ* with his model. This is done in Fig. 18, where the contractile element's shortening velocity is plotted vs muscle force at "normal resting length." The transformations in the Appendix were used to convert angular velocity and torque to translational velocity and force, respectively. The main difference is in the nonlinearity of those curves; Hill's have a more pronounced curvature. They both display a similar decaying shape for the velocities involved, however, with the exception of the region close to equilibrium which is attributed to the very large changes in velocities observed in our data near the equilibrium region.

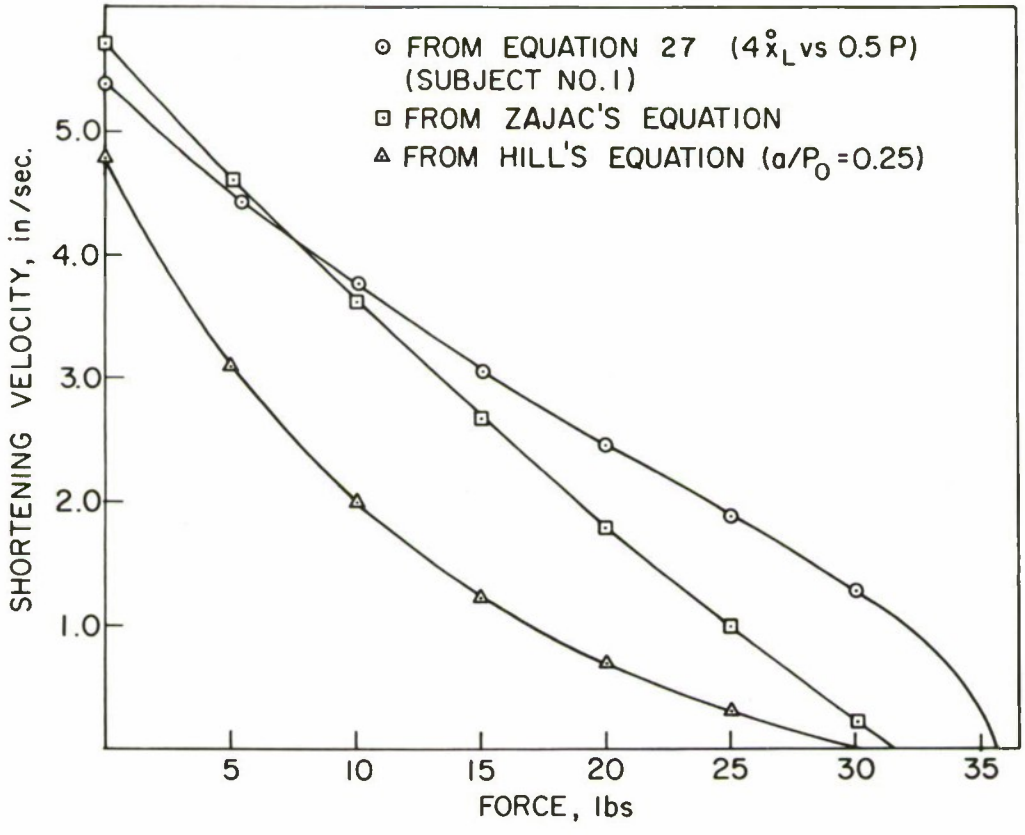


Fig. 18. VELOCITY-FORCE CURVES FOR CONTRACTILE ELEMENTS AT NORMAL RESTING LENGTHS.

Differences are expected in these results on the basis of the differences in experimental conditions (in vitro vs in situ) and on the fact that Hill performed his experiment on frog muscles and not on human skeletal muscle. The long-discovered discrepancies between model predictions and actual data collected for frog muscles under similar environmental conditions have motivated many investigators to re-examine the theoretical results for muscle behavior. Jewell and Wilkie [Ref. 11] tested the model extensively and reported a "clear" discrepancy between the theoretical prediction and their experimentally obtained velocity-force curve. Carlson [Ref. 12] also showed that Hill's equation could not describe his data at the longer muscular lengths.

Fenn and Marsh [Ref. 13] and more recently Zajac [Ref. 4] studied the shortening characteristics of cat muscle. Zajac, in particular, used isotonic shortening data to model the kinematic properties of the cat's medial gastrocnemius muscle in vivo. His development was not restricted to one specific muscle length but elegantly reflected the dependence of the contractile-element velocity on both muscle force and length. His results are given by

$$\dot{x}_C = V_o \exp(-P/a_v)\left[\left(\frac{|L_o - x_L + k_x P|}{L_o - L_S}\right)^c - 1\right] \tag{36}$$

where

$\dot{x}_C$ = contractile-element velocity

x_L = muscle length

P = muscle force

and the other parameters are constants, for a given muscle, obtained empirically from his experimental data. A great deal of similarity can be observed between Zajac's model and the one derived in this investigation. The notable differences, however, are the following:

(1) The shape of the velocity-length curves for a constant force are defined by the power constant c for each muscle, where c is the slope obtained empirically by plotting velocity vs length on log-log scales. Such a description provides great generality and can closely match the elliptical shapes characterized in our model, excluding the region near equilibrium.

(2) If the concept of a velocity-length curve for a nominal force is considered, which is rescaled and translated so as to describe the data for forces other than the nominal, then it is observed that no scaling along the muscle-length axis is implied in Zajac's model.

These differences are illustrated in the following general representations:

$$\dot{\theta}_C = h_T(T_A) \cdot f_T\left[\theta_L \ell_T(T_A) - g_T(T_A)\right] \tag{37}$$

$$\dot{x}_C = h_p(P) \cdot f_p\left[x_L - g_p(P)\right] \tag{38}$$

where (37) represents the model obtained in this study, and (38) is the model obtained by Zajac for cat muscle. Difference (1) is the dissimilarity in the functions $f_T(\cdot)$ and $f_p(\cdot)$, where the former is an elliptical function and the latter is described by the power constant c. Difference (2) appears as the lack of a counterpart for the position-gain or scaling function $\ell_T(T_A)$ in Zajac's equation. This has the surprising implication that all the velocity-length curves obtained for cat muscle under different loads have identical shapes near the equilibrium length, and is illustrated by Fig. 19, where the curves for the larger forces are translated to shorter muscle lengths to indicate the identical shapes (overlapping curves) that they would have in the neighborhood of equilibrium. This was not the case for our experiments on human muscles.

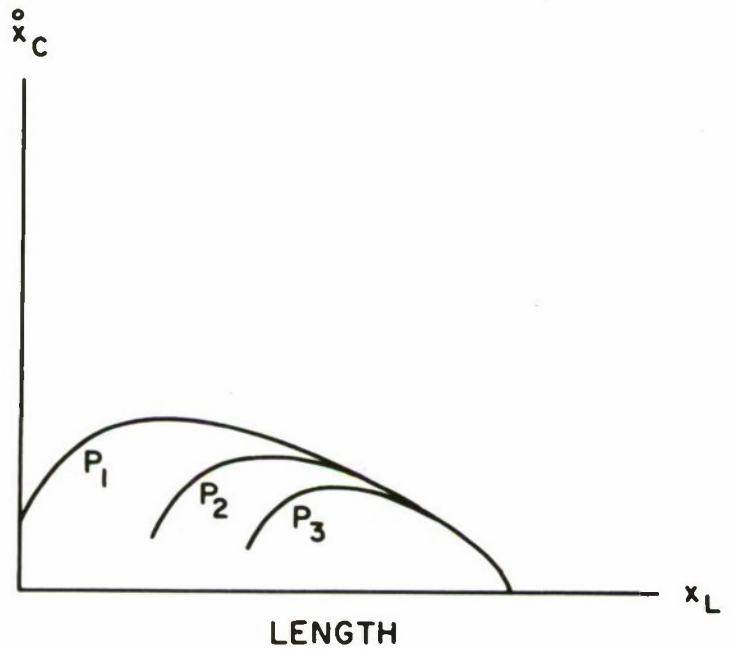


Fig. 19. OVERLAPPING PHASE-PLANE PLOTS FOR DIFFERENT LOADS NEAR MUSCLE EQUILIBRIUM LENGTH, AS IMPLIED BY ZAJAC'S EQUATION. $P_3 > P_2 > P_1$. x_L = distance shortened. Curves for P_2 and P_3 translated to shorter muscle length.

The velocity-force curve for Zajac's equation is also included in Fig. 18 for comparison with our results and Hill's; its shape resembles the velocity-force curve derived from this investigation except in the neighborhood of the equilibrium position, as anticipated. It is also observed from these curves that, although significantly larger loads could be moved by human muscles, the shortening velocities of frog and cat muscles are two to four times faster.

An angular velocity-muscle torque plot at normal resting length is presented in Fig. 20 for completeness; also included is a maximum velocity-torque curve, given by

$$V_{max} = h(T_A) = Ue^{-aT_A} \tag{39}$$

with the observation that the maximum shortening velocity occurs at different muscle lengths for different loads. The muscular lengths at which maximum velocity occurs are

$$\theta_L\Big|_{V_{max}} = g(T_A) = b - mT_A \tag{40}$$

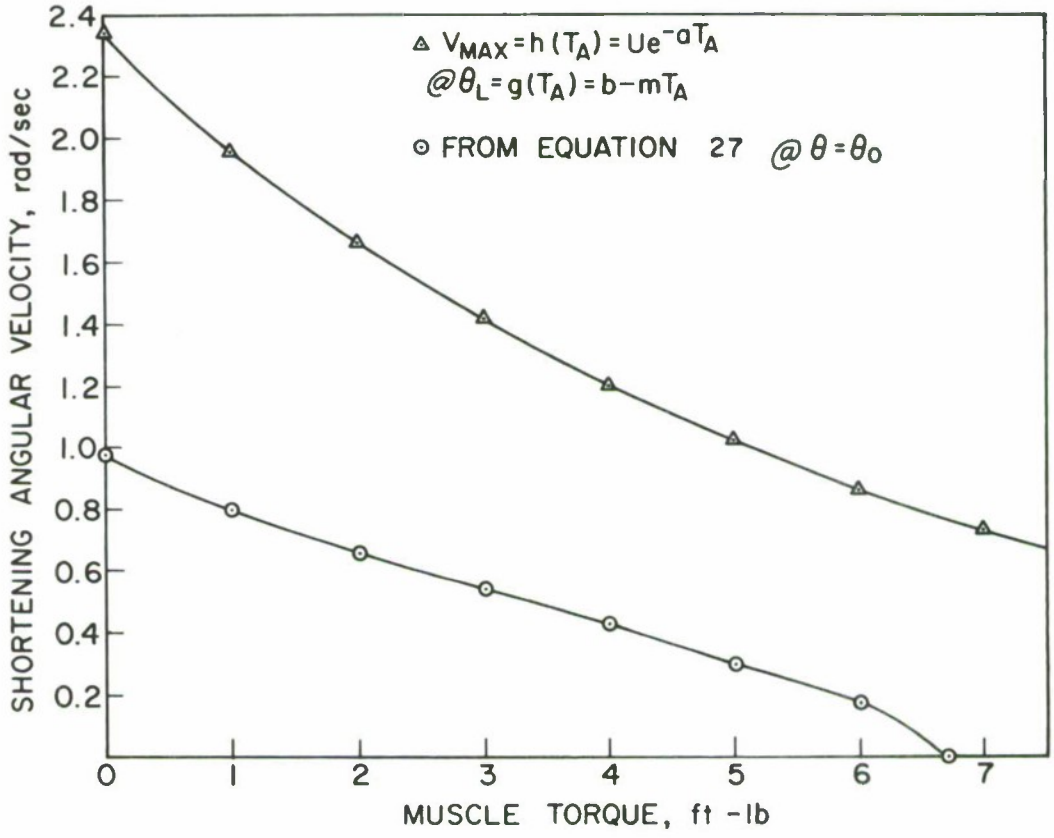


Fig. 20. ANGULAR-VELOCITY VS TORQUE CURVES FOR CONTRACTILE ELEMENT. Subject 1.

Appendix

RELATIONSHIPS BETWEEN TRANSLATIONAL AND ROTATIONAL COORDINATES AND COMPUTATION OF MUSCLE-GROUP TORQUE ABOUT THE ELBOW JOINT

If a set of unit vectors $\underline{i}$, $\underline{j}$, $\underline{k}$ (considered fixed in the muscle) is defined as shown in Fig. 21 and if

$x \equiv$ length of muscle group

$d \equiv$ distance between muscle-group attachment point at forearm and elbow joint

$\ell \equiv$ distance between muscle-group attachment point at shoulder and elbow joint

$\underline{P} \equiv$ tension vector in muscle

$\underline{T_A} \equiv$ muscle-group torque vector about elbow

$\underline{q} \equiv$ moment arm vector of muscle force about elbow joint

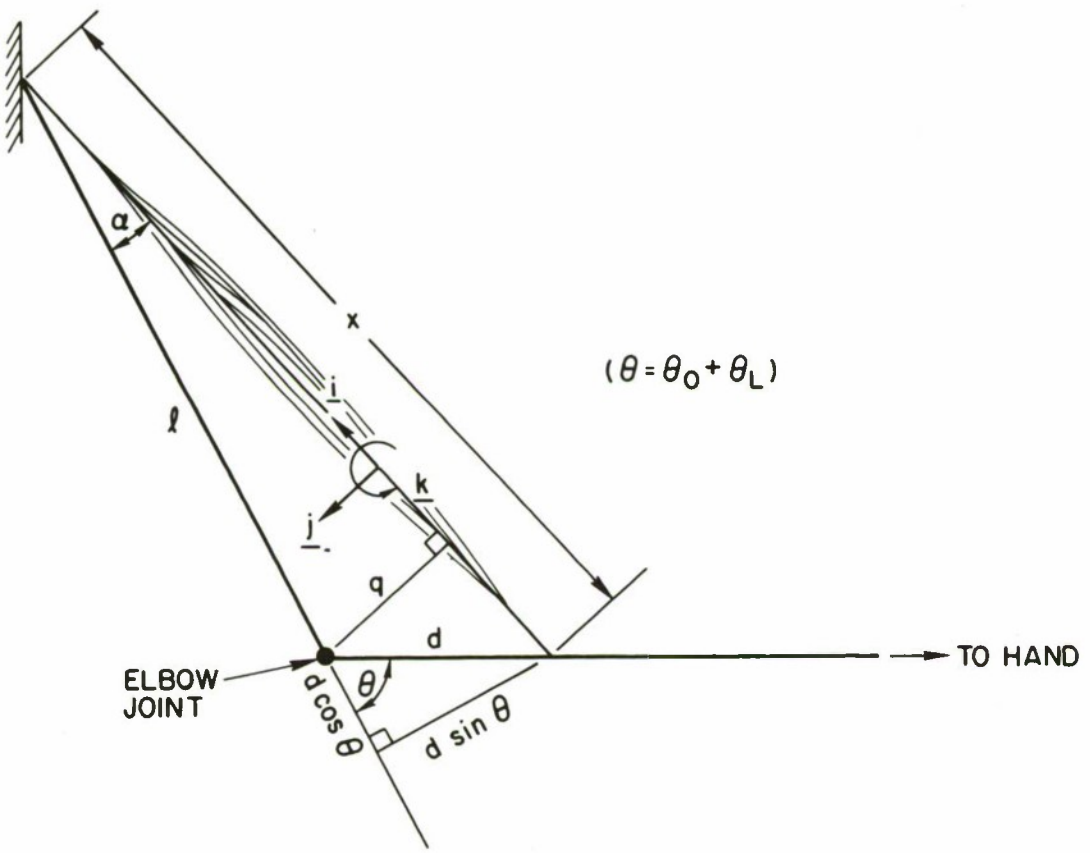


Fig. 21. GEOMETRIC REPRESENTATION OF ARM/MUSCLE-GROUP SYSTEM.

and angles α and θ are defined as indicated in Fig. 21, then the following relationships can be written:

$$x^2 = (\ell + d \cos \theta)^2 + (d \sin \theta)^2$$

or

$$x = (\ell^2 + d^2 + 2\ell d \cos \theta)^{1/2} \qquad (A.1)$$

and, for the definitions of x_C, x_L, θ_0, θ_C, and θ_L,

$$\underline{x_C} = \left(\ell^2 + d^2 + 2\ell d \cos \theta_0\right)^{1/2} - \left[\ell^2 + d^2 + 2\ell d \cos (\theta_0 + \theta_C)\right]^{1/2} \underline{i} \quad (A.2)$$

and

$$\underline{\dot{x}_C} = \frac{\ell d \sin (\theta_0 + \theta_C)}{\left[\ell^2 + d^2 + 2\ell d \cos (\theta_0 + \theta_C)\right]^{1/2}} \dot{\theta}_C \underline{i} \quad (A.3)$$

Similarly,

$$\underline{x_L} = \left(\ell^2 + d^2 + 2\ell d \cos \theta_0\right)^{1/2} - \left[\ell^2 + d^2 + 2\ell d \cos (\theta_0 + \theta_L)\right]^{1/2} \underline{i} \quad (A.4)$$

$$\underline{\dot{x}_L} = \frac{\ell d \sin (\theta_0 + \theta_L)}{\left[\ell^2 + d^2 + 2\ell d \cos (\theta_0 + \theta_L)\right]^{1/2}} \dot{\theta}_L \underline{i} \quad (A.5)$$

Now,

$$\underline{q} = -\ell \sin \alpha \underline{j} = -\frac{\ell d \sin \theta}{x} \underline{j}$$

and use of Eq. (A.1) yields

$$\underline{q} = \frac{-\ell d \sin \theta}{\left(\ell^2 + d^2 + 2\ell d \cos \theta\right)^{1/2}} \underline{j} \quad (A.6)$$

By definition, the torque of the muscle group about the elbow is

$$\underline{T_A} = \underline{q} \times \underline{P}$$

and because $\underline{P} = P\underline{i}$ and $\theta = \theta_0 + \theta_L$,

$$\underline{T_A} = \frac{P\ell d \sin (\theta_0 + \theta_L)}{\left[\ell^2 + d^2 + 2\ell d \cos (\theta_0 + \theta_L)\right]^{1/2}} \underline{k} \quad (A.7)$$

By noting that θ_0 is a constant and making use of the above results, Eqs. (A.2), (A.3), (A.4), (A.5), and (A.7) can be represented in general forms as

$$x_C = h_1(\theta_C) \quad (A.8)$$

$$\dot{x}_C = \dot{\theta}_C h_2(\theta_C) \quad (A.9)$$

$$x_L = h_3(\theta_L) \tag{A.10}$$

$$\dot{x}_L = \dot{\theta}_L h_4(\theta_L) \tag{A.11}$$

and

$$P = T_A h_5(\theta_L) \tag{A.12}$$

where

$$h_5(\theta_L) = \frac{\left[\ell^2 + d^2 + 2\ell d \cos(\theta_0 + \theta_L)\right]^{1/2}}{\ell d \sin(\theta_0 + \theta_L)} \tag{A.13}$$

REFERENCES

1. A. V. Hill, "The Heat of Shortening and the Dynamic Constants of Muscle," Proc. Roy. Soc., B, 126, 1938, pp. 136-195.

2. D. T. McRuer and E. Krendel, "Dynamic Response of Human Operators," WADC TR 56-524, 1957. DDC No. AD 110 693

3. A. Levin and J. Wyman, "The Viscous Elastic Properties of Muscle," Proc. Roy. Soc., B, 101, 1927, pp. 218-243.

4. F. E. Zajac, "The Mathematical Formulation of the Kinematic Properties of Muscle Derived from an Experimental Investigation," Ph.D. dissertation, Stanford University, Stanford, Calif., 1968.

5. L. Vodovnik, W. J. Crochetiere, and J. B. Reswick, "Control of a Skeletal Joint by Electrical Stimulation of Antagonists," Med. and Biol. Engineering, 5, 1967, pp. 97-109.

6. L. J. Leifer, "Characterization of Single Muscle Fiber Discharge during Voluntary Isometric Contraction of the Biceps Brachii Muscle in Man," Ph.D. dissertation, Stanford University, Stanford, Calif., 1969.

7. J. M. Ritchie and D. R. Wilkie, "The Dynamics of Muscular Contraction," J. Physiol., 143, 1958, pp. 104-113.

8. A. V. Hill, "Water and Heat in a Muscle Twitch," Proc. Roy. Soc., B., 136, 1949, pp. 220-228.

9. A. V. Hill, "The Abrupt Transition from Rest to Activity in Muscle," Proc. Roy. Soc., B., 136, 1949, pp. 399-420.

10. B. C. Abbott and D. R. Wilkie, "The Relation between Velocity of Shortening and the Tension-Length Curve of Skeletal Muscle," J. Physiol., 117, 1953, pp. 77-86.

11. B. R. Jewell and D. R. Wilkie, "An Analysis of the Mechanical Components in Frog's Striated Muscle," J. Physiol., 143, 1958, pp. 515-540.

12. F. D. Carlson, "Kinematic Studies on Mechanical Properties of Muscle," in Tissue Elasticity (J. W. Remington, ed.), Am. Physiol. Soc., Washington, D.C., 1957.

13. W. O. Fenn and B. S. Marsh, "Muscular Force at Different Speeds of Shortening," J. Physiol., 85, 1935, pp. 277-297.

14. E. V. Evarts, "Relation of Pyramidal Tract Activity to Force Exerted during Voluntary Movement," J. Neurophysiol., 1967.

15. L. D. Partridge, "Signal-Handling Characteristics of Load-Moving Skeletal Muscle," Am. J. Physiol., 5, 210, 1966, pp. 1178-1191.

16. R. Alter, "Bioelectric Control of Prostheses," MIT Research Lab. of Electronics, TR No. 446, Cambridge, Mass., 1966.

17. T. G. Ruch, "Motor Systems," in Handbook of Experimental Psychology, (S. S. Stevens, ed.), John Wiley & Sons, New York, 1951, pp. 154-208.

18. D. T. McRuer et al, "Human Pilot Dynamics in Compensatory Systems--Theory, Models, and Experiments with Controlled Element and Forcing Function Variations," AFFDL, TR No. 65-15, 1965. DDC No. AD 470 337

19. D. T. McRuer, R. E. Magdaleno, and G. P. Moore, "A Neuromuscular Actuation System," USC-NASA Conf. on Manual Control, 1967.

20. W. J. Crochetiere, L. Vodovnik, and J. B. Reswick, "Electrical Stimulation of Muscle--A Study of Muscle as an Actuator," Med. and Biol. Engineering, 5, 1967, pp. 111-125.

21. D. T. Greenwood, Principles of Dynamics, Prentice-Hall, Inc., New Jersey, 1965.

22. B. Katz, "The Relation between Force and Speed in Muscular Contraction," J. Physiol., 96, 1939, pp. 45-64.

23. B. Bigland and O. C. J. Lippold, "The Relation between Force Velocity, and Integrated Electrical Activity in Human Muscles," J. Physiol., 123, 1954, pp. 214-224.

24. D. R. Wilkie, "The Mechanical Properties of Muscle," Brit. Med. Bull., 3, 12, 1958, pp. 177-182.

25. X. Aubert, "Le Couplage Energetique de la Contraction Musculaire," thèse d' agrégation, Brussels, 1956.

26. C. R. Dorf, Modern Control Systems, Addison-Wesley Series in Electrical Engineering, 1967.

27. J. E. Gibson, Nonlinear Automatic Control, McGraw-Hill Book Co., New York, 1963.

28. B. Katz, Nerve, Muscle, and Synapse, McGraw-Hill Series in New Biology, 1966.

Sensory Feedback in Human Posture Control*

Lewis M. Nashner and Jacob L. Meiry
Massachusetts Institute of Technology

1. Introduction

Current models for physiological components and a series of experiments on human subjects form the basis for a multiloop control model which describes how a human uses multiple feedback sensors to control his orientation. Particular emphasis is placed on defining functional interfaces between the feedback sensors and postural responses. Because of the inherent complexities within the posture control system, analysis is simplified by considering only control of forward and backward rotational motions about the ankle joints during quiet standing tasks.

Because of the inherent complexity of the posture control system, analysis is simplified by considering only control of forward and backward rotational motions of the body about the ankle joint, hereafter termed body sway or body angle motion. Body sway motion represents the critical mode in control of posture because of the inherently unstable "inverted pendulum" characteristics of the body. Therefore, this simplification is justified. Relative motion between upper body segments is of considerably less consequence during quiet standing and may be neglected here. Hence, the goal of the postural control system in this simple task is to assess the current status of body sway motion and generate appropriate ankle reaction torques to maintain stability.

The research is divided into three segments. First, a general posture control model is assembled, given current models for motor and sensory components and the general properties of neural processing described in the literature. This general model then forms the basis for a series of experiments. Finally, experimental observations are combined with the general physiological model to develop specific models for the sensory-motor interface.

In both the design of the experiments and development of the final models, the highly adaptable and non-stationary characteristics of the system are recognized. Experiments use transient disturbances which probe the states of the system at specific instances in time or during very short periods in which characteristics remain relatively constant. The models also define transient rather than continuous control processes.

Sensory deprivation techniques are employed during experiments to isolate characteristics of the individual sensory feedback modes. These methods allow observation of a number of control strategies and enable a more complete determination of the range of adaptability of the various sensory feedback modes.

*This research is supported by NASA Grant NGR-22-009-156.

2. Reflex Control of Posture

Two direct mechanical mechanisms resist length changes in flexor and extensor muscles. During constant stimulation, muscle tensions vary in proportion to small length changes about a median length (13, 17). Rate of stretch also affects tension of the muscle under constant stimulation (5). Both the length-change and velocity induced tensions act to resist overall change in muscle length. It should be noted that the sensitivity of both of these mechanisms varies in response to the activation level of the muscle itself.

In addition to the "mechanical" mechanisms varying tension to resist muscle length change, there is active resistance initiated by muscle spindle reflexes.

Stretching skeletal muscle produces a corresponding stretch of the muscle spindle fibers interspersed throughout the muscle. Stretching a muscle spindle increases its afferent discharge, which then acts through the alpha motoneuron pool to contract skeletal muscle fibers in the region of the stretched spindle. Increase of muscle tension is achieved by a combination of the recruitment of additional motor units and increased activation of units already firing.

Higher center activity influences the reflex feedback functions in the following ways:

1. Muscle "mechanical" properties are dependent on the level of muscle activation (10,13).
2. Muscle spindle feedback "gains" are controlled by higher center commands (8).
3. Higher centers may act on the alpha-motoneuron pool, exercizing direct control over the activation level of the muscle.

In describing control mechanisms of the reflex arc, physiologists have recognized its likeness to a position sevo-mechanism. Only recently, however, have investigators integrated functional descriptions of the individual physiological components into models of the reflex control system (7, 9, 13).

These studies demonstrate the validity of mathematical reflex control models. Most importantly, they show that linearized component models and equivalent bilateral muscle models can predict postural responses. Further investigations of posture control mechanisms described in this paper begin with these premises.

The reflex control mechanism, including proprioception feedback pathways and higher center inputs, is illustrated in Figure 1. Components include postural muscles responsive to both active neural stimulation and to mechanical stretch and position feedback via the muscle spindles which respond in proportion to length and rate of stretch. The higher centers excercise independent control of both phase and position feedback gains.

Reflex Responses During Quiet Standing

The following is a description of the experiment conducted to determine specific values of the parameters for the ankle stretch reflex control loop. The reflex response parameters obtained are then formulated into a model which quantifies the contribution of the ankle

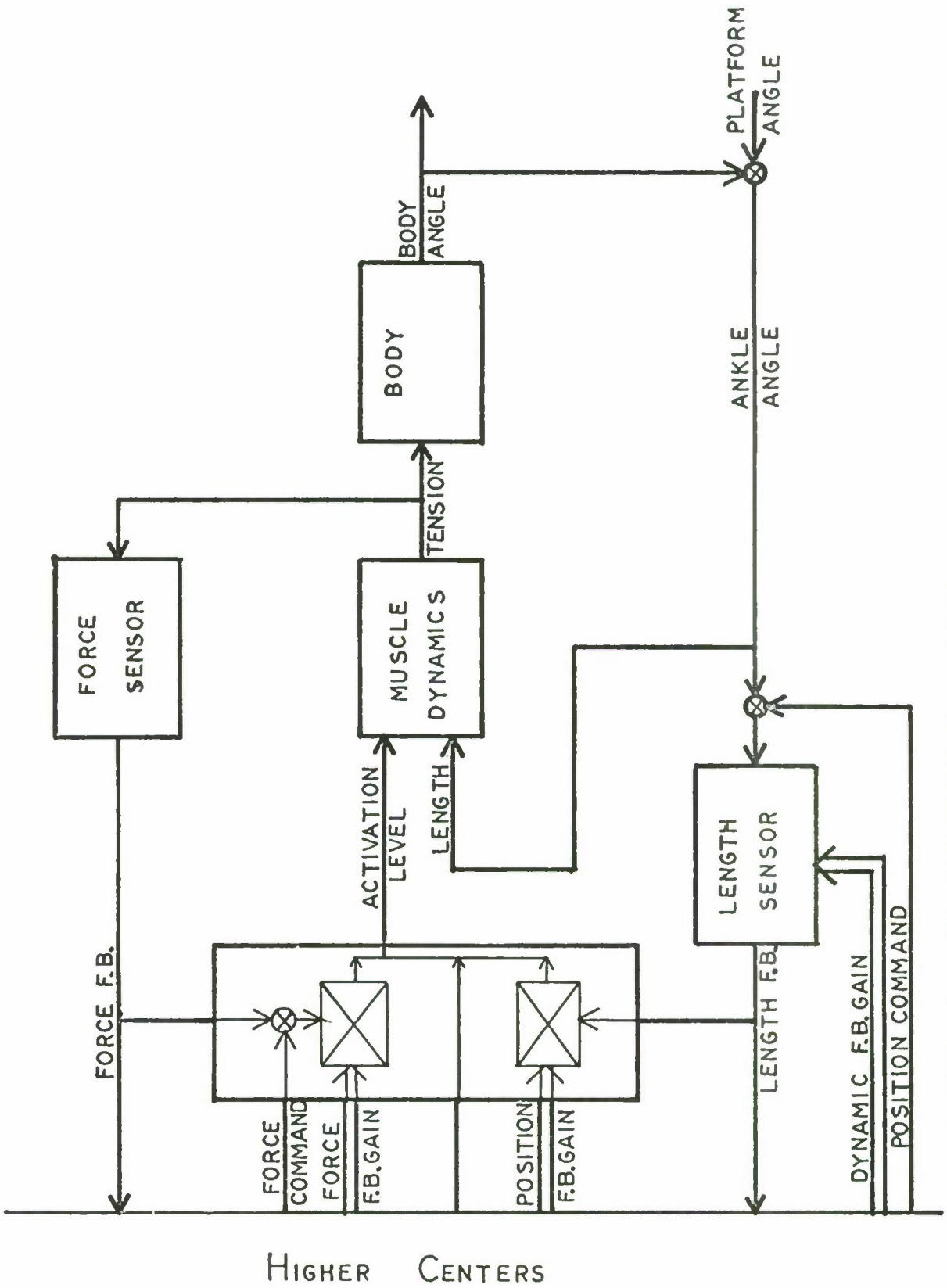


FIGURE 1 THE REFLEX CONTROL SYSTEM

reflex control loop to overall postural stability, providing the basis for further experiments exploring the strategy of reflex parameter adaptation to changing control conditions.

The subject stands relaxed on the experimental platform, arms folded above the waist and feet ten to twelve inches apart. The platform detects the subject's reaction torques and his sway angle about the ankle joint. A hydraulic position servo allows rotations of the platform about an axis colinear with that of the subject's ankle joints. The ankle stretch reflex is excited by small steps of the platform.

During each test run the subject is asked to stand relaxed with his eyes opened. To prevent fatigue, each test run is limited to 16 step samples, or approximately 3 minutes. Seven test runs are conducted for each of five step sizes: 1/10, 1/4, 1/2, and 1 1/2 degrees. A test run consists of one step size only. The ordering of step size test runs is random. Responses are characterized by a peak within 80-125 milliseconds, a compensatory response peak at 400-800 milliseconds, then gradual return to equilibrium.

The reflex response amplitude, defined as the maximum response occurring within 80 to 125 milliseconds after initiation of the step disturbance, is determined for each step response. Reflex gain is defined as response amplitude divided by the ankle step size. The composite average gain as a function of step size is shown in Figure 2. Variation among subjects is statistically insignificant, $p > 0.1$.

To examine the later compensatory phase, average responses of each subject to the 1/4 and 1/2 degree steps are computed. These response groups are further subdivided according to initial reflex gain into three levels. Variations among the three subjects are statistically insignificant, $p > 0.1$. As an example, Figure 3 shows the response averages for one test case.

A significant feature of the reflex control loop is the large increase in gain for disturbances of very small amplitude. While gains for steps of 1/4 to 1/2 degrees are considerably below that necessary for postural stability (2 ft-lb/degree versus 7 ft-lb/°) suggests that reflex control alone may fully stabilize the body for disturbances below 1/10°. This is substantiated by current physiological evidence.

Evidence of large gain increases in muscle responses to small length changes is found in the work of Hill (6) and Joyce et al (11). Similar gain increases in reflex responses are shown by Brown et al (1) and Matthews (14).

Recordings of ankle torque and body angle responses during quiet standing indicate that the reflex loop fully stabilizes very small amplitude body sway motions. All subjects exhibit frequent periods of five or more seconds duration in which no motion within measurement resolution ($T = \pm 0.10$ ft-lb θ_B ± 0.02 degrees) can be observed. All higher center feedback controls are limited by thresholds considerably larger than the measurement resolutions. (Results, Sections 3 and 4). Therefore, reflex feedback control must be responsible for these stable periods.

On the basis of forementioned physiological and experimental evidence, a reflex gain function is defined in which both muscular and reflex responses show gain increases due to stiction for deflections less than 0.15°. Reflex feedback gain is sufficient for complete postural stability for deflections less than 0.05°.

The compensatory response is initiated by higher center commands. Observed delay time, about 200 milliseconds, agrees with physiological evidence for higher center response delay; it also agrees with vestibular response delays measured by the author.

Figure 4 shows the compensatory response natural frequency as a function of the initial reflex disturbance amplitude. Included in the figure for comparison is the response characteristics as a function of gain for the following system: an inverted pendulum stabilized with a rate compensated feedback, with system dynamics roughly equivalent to those of the body reflex control system. Results suggest that higher centers enhance the total reflex gain during a transient disturbance in proportion to the perceived disturbance amplitude. Reports of several investigators also indicate reflex gain control as a likely higher center control mechanism.

Kim and Partridge (12) show that total reflex gain (Δ tension/ Δ length) in cat soleus muscle is enhanced by factors of two to three times when the vestibular nerve is stimulated continuously. They note that neck rotations also enhance reflex gain. In a further study, Partridge and Kim (16) find that isometric tension varies less than 1% of maximum muscle tension during sinusoidally modulated vestibular stimulation.

Gernandt et al. (4) have shown that vestibular stimulation in anesthetized cats strongly effects the dorsal to ventral root response amplitude at the cervical and lumbrosacral levels of the spine.

In summary, vestibular and exteroceptive stimulation is shown to enhance reflex responses, establishing this mechanism as a probable method for higher center control of postural disturbances. The fact that vestibular responses do not act as length commands gives further support to this conclusion. Two physiological mechanisms for reflex enhancement are proposed: multiplicative increase in the dorsal to ventral reflex gain, and control of the tension levels of muscle. Functionally the mechanisms are equivalent.

The strategy of gain enhancement shows an excellent combination of higher and lower levels of sensory information. The higher center gain command involves no complex processing and can be effected with a minimum of delay. Additional corrective computation takes place at the spinal level through enhanced resistance of body displacements. This mechanism is a crude one, but in most situations a good "first guess." More complex higher level mechanisms requiring longer processing time may intervene somewhat later to fine-tune the initial crude corrections.

The Role of Ankle Reflexes in Postural Stability

The ankle stretch reflex feedback helps stabilize body sway motion. In many circumstances, however, independent reflex control cannot perform this function. On non-rigid surfaces, inertial information is lost and the higher center sensory loops (eyes and vestibular organs) must mediate or override the reflex responses to provide postural stability.

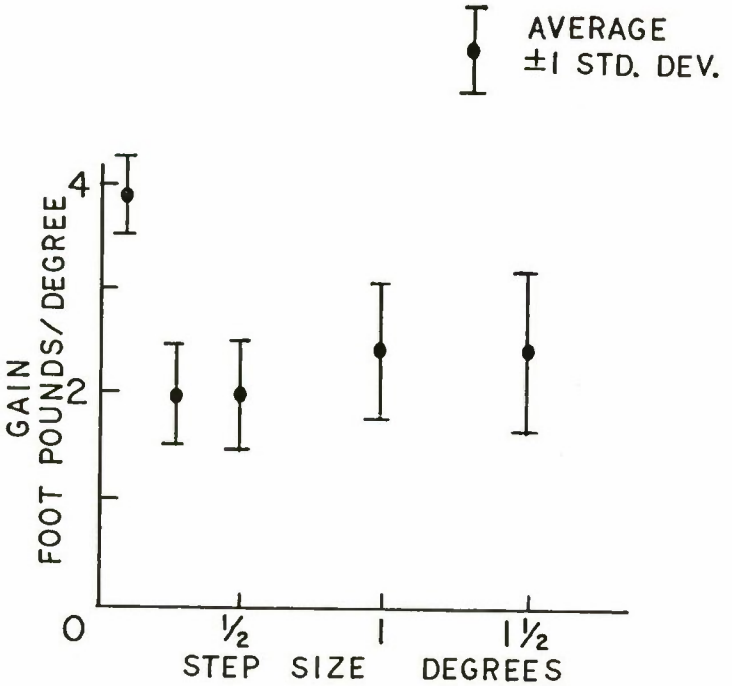


FIGURE 2 AVERAGE ANKLE REFLEX GAIN AS A FUNCTION OF STEP SIZE

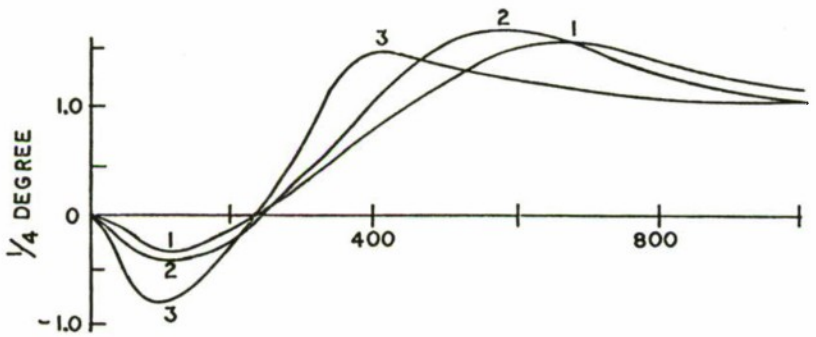


FIGURE 3 AVERAGE ANKLE REFLEX RESPONSES GROUPED ACCORDING TO GAIN FOR 1/4 DEGREE STEPS FLEXING THE ANKLE

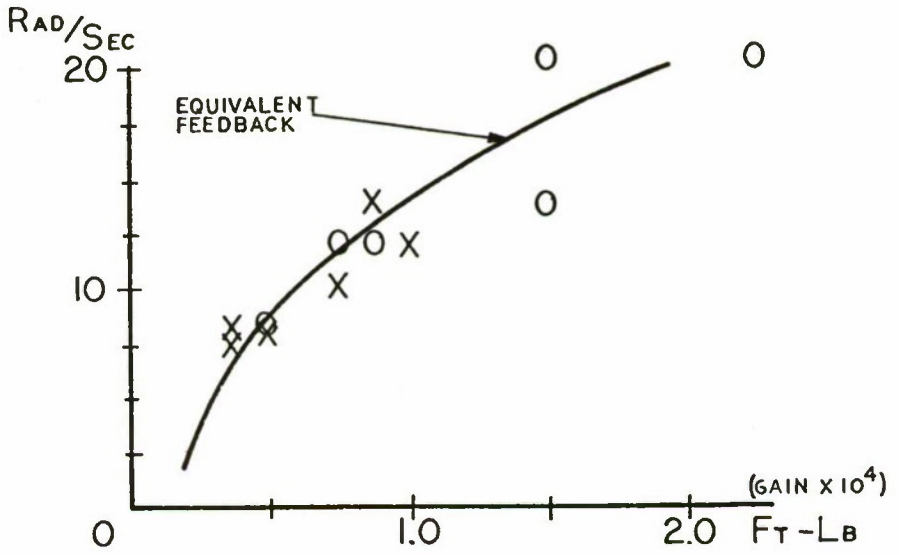


FIGURE 4 NATURAL FREQUENCY AS A FUNCTION OF REFLEX AMPLITUDE, COMPENSATORY RESPONSE

The following experiments explore fully the relationship of reflex and higher center control strategy, illustrating reflex gain during a variety of circumstances which alter the effectiveness of the reflex control loop. Test conditions are designed to determine the effects of the following on feedback parameters:

1. visual feedback: enhancement and elimination
2. conscious posture set
3. random ankle angle disturbances
4. elimination of ankle position feedback

All subjects clearly demonstrate adaptation of the reflex gain in response to changes in posture control conditions. Subjects are able to increase reflex gain when asked to exert special effort to minimize sway. They all reported, however, that this condition was quite uncomfortable if maintained for more than a few minutes.

When given a high sensitivity visual display of body angle, subjects' reflex gain decrease only slightly compared to normal gain. Subjects reported that this condition too was tiring after several minutes.

When small random disturbances are introduced, reflex gain decreases to nearly one-half the normal gain.

The median gain is nearly zero when the reflex control loop is fully suppressed by opening the position feedback loop.

Elimination of visual feedback has little effect on reflex control strategy under any of the tested conditions.

The availability of the supporting surface for body and inertial reference information determines the reflex gain setting. Reflex gain is reduced as the reflex control loop becomes a less effective mode of stabilization. Changes in the state of higher center feedback sensors seem to have no effect on the strategy for reflex gain setting.

3. Vestibular Control of Posture

The vestibular organs, the utricles, the saccules, and the semicircular canals are the prime motion sensors in the human. Vestibular cues contribute important sensory information for the regulation of posture.

Characteristics of the vestibular system are well documented (15, 19, 20). A complete description can be found in the references listed above; a review of the details pertinent to posture control is given here.

Angular acceleration in three dimensional space is sensed by three approximately orthogonal semicircular canals in each inner ear. The utricle otoliths, one in each inner ear, are multi-dimensional linear accelerometers. They sense specific forces (linear acceleration plus gravity) in a plane rotated 30 degrees with reference to the horizontal plane of the head. Hence, combined canal receptors sense all relevant angular motions of the body; utricle otoliths sense the summation of all linear and gravitational forces.

The canals are heavily damped accelerometers, with perceived output corresponding to angular velocity. The threshold of perception of angular velocity is heavily dependent on the sensory mode in which it is measured. A discussion of this point is included in the following paragraphs. The utricle otoliths are also heavily damped accelerometers, sensing both tilt angle and linear velocity.

Angular acceleration threshold values for the semicircular canals reported in the literature vary over a wide range. Meiry (15) estimates a linear acceleration threshold of 0.005g, equivalent to a 0.30° threshold for subjective orientation. Clark and Stewart (2) find detection of steady state angular accelerations with the oculogyral illusion at values ranging from 0.04 to 0.28°/sec^2. The illusion does not involve a direct sensation of movement, rather the subject, seated in a dark enclosure during angular acceleration, perceives movement of a light spot which is actually fixed with respect to him. The illusion of movement is believed to be caused by the action of canal output responses on the visual system. Their results show that canal responses influence visual sensation at levels of angular acceleration significantly below the levels which can be detected subjectively.

Vestibular Stimulation by Body Angle Motion

Angular acceleration of a rigid human body about the ankle joints is directly equivalent to the accelerating input to the pitch axis semicircular canals. Since the posture control model to be developed here considers only quiet standing during which the upper body remains relatively rigid, equivalence of body angle and pitch angular accelerations may be assumed.

Stimulation of the utricle otolith organs by body sway motions is considerably more complex since both gravitational and linear acceleration reaction forces act as inputs. During uncontrolled sway divergence of the body, gravitational (f_g) and tangential acceleration reaction force (f_t) act in opposite directions, Figure 5.

$$I_B \ddot{\theta}_B = m_B gh_{cg} \theta_B + T_A \qquad 3.1$$

The net reaction force on the otolith is:

$$f_o = \theta_B \left(g - \frac{m_B g h_{cg} d}{I_B} \right) \qquad 3.2$$

Because the utricle otolith organs are located well above the center of mass of the body, the net reaction force on the otolith is negative during free-fall sway divergence of the body, (i.e. it is opposite to the force on the otolith due to gravity alone). If the body diverges with partial resistence from postural responses, the net force acting on the otolith may be negative, zero, or positive depending on the intensity of postural resistence. Recalling that reflex responses oppose body sway with a torque roughly proportional to angular deflection of the ankle joint, the following relations demonstrate the ambiguity of the sign of the linear motion sensor:

$$T_A = -K \theta_B \qquad 3.3$$

$$f_o = \theta_B \left(g - \frac{m_B g h_{cg} d_o}{I_B} + \frac{K d_o}{I_B} \right) \qquad 3.4$$

Using body parameter values (18) we find that:

$f_0 > 0$ if K > 4 ft-lb/degree

$f_0 < 0$ if K < 4 ft-lb/degree

During stabilization postural responses, tangential and gravitational forces on the otolith are always of the same sign.

Postural Response to Vestibular Stimulation

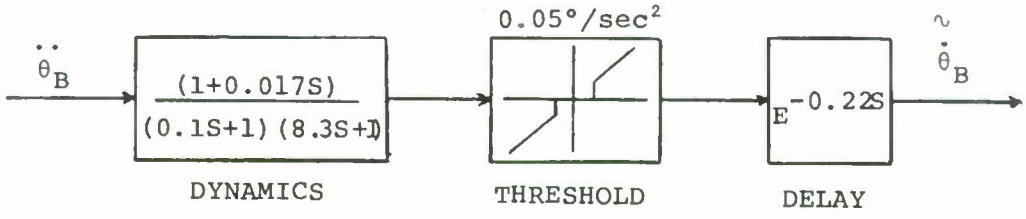
The subject stands relaxed on the experimental platform. Sway is induced using small forward or backward platform displacements at constant velocity, introduced by the experimenter when the subject shows no movement. The platform disturbance induces a step change in body angular velocity.

Elimination of all other modes of sensory feedback is necessary to insure that postural responses to the induced sway are vestibular in origin. To remove reflex responses and extereoceptive cues, the platform is rotated to track the motions of the body; thus, maintaining an ankle angle of zero. Nulling ankle angle during body angle motion effectively opens the reflex position feedback loop but does not interfere with the subject's ability to generate isometric ankle control torques. Removal of reflex feedback also eliminates any advanced extereoceptive cues. Subjects are tested with eyes open and eyes closed to determine effects of visual cues on vestibular detection of body sway.

Figure 6 shows typical postural responses to rapid and to slow induced body motion stimuli. Responses are characterized by an initial period during which the body angle begins to increase while the ankle torque remains unchanged. When the motion is detected by the vestibular sensors, ankle reaction torque increases rapidly, returning the body to a stable position. Threshold of the vestibular feedback sensors is defined in terms of the time after initiation of motion at which the ankle torque level has increased 0.25 foot pounds above its initial level.

Conclusions - The Vestibular Model

Threshold characteristics for forward and for backward sway are identical. Variations among the subjects and among the test conditions are statistically insignificant, $p > 0.1$ for all samples. Composite threshold functions for the three test conditions are compared to the semicircular canal model threshold characteristics in Figure 7. The following semicircular canal model predicts the observed threshold characteristics over the complete range of induced sway rates.



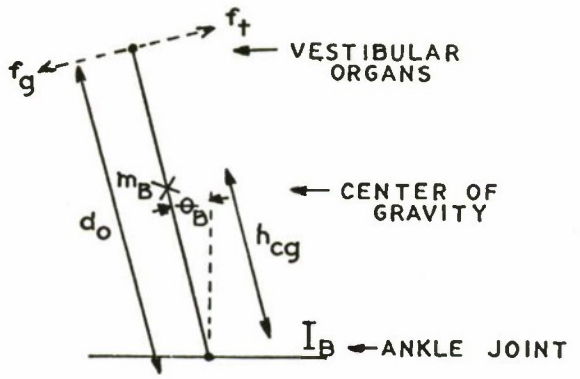


FIGURE 5 THE BODY AND ITS SWAY DYNAMICS

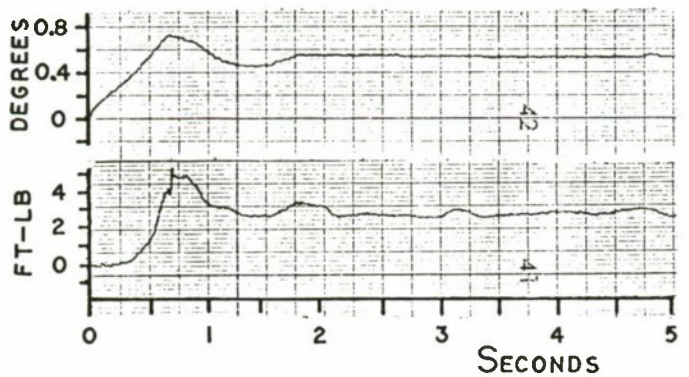


FIGURE 6 A TYPICAL POSTURAL RESPONSE TO INDUCED SWAY MOTION

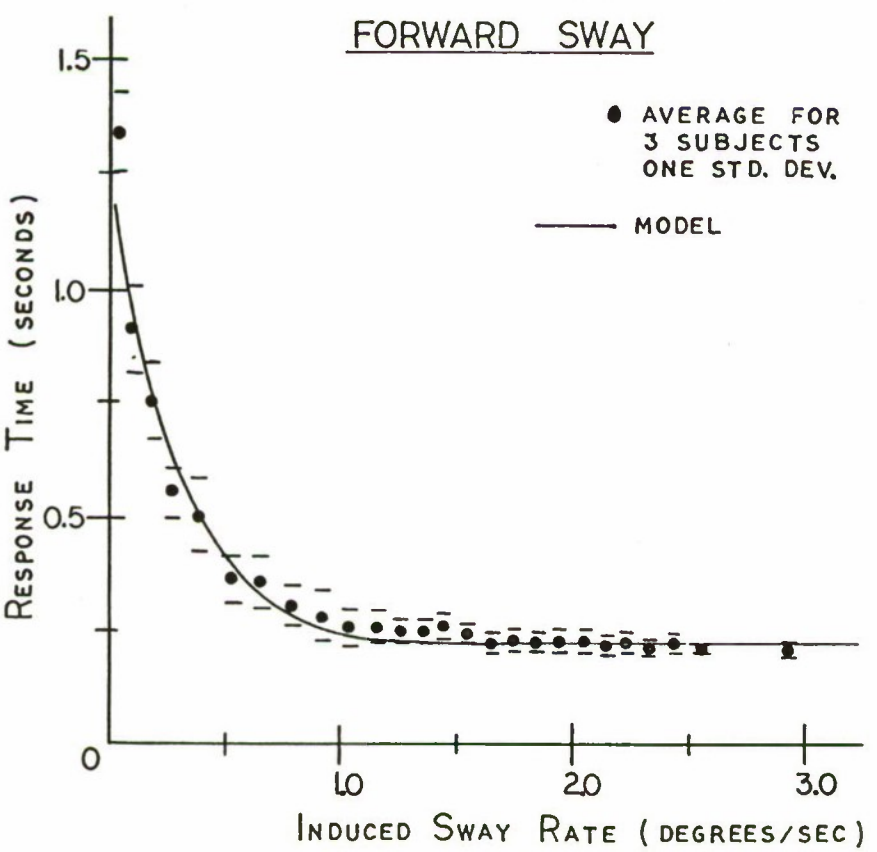


FIGURE 7 VESTIBULAR THRESHOLD RESPONSE CHARACTERISTICS TO INDUCED SWAY MOTION

Linear dynamic characteristics of the semicircular canal model compare closely with Meiry's canal model. The addition of a very small lead term is necessary to predict the minimum response time characteristics for large impulsive inputs. This term has no effect on threshold response properties within the dynamic range of normal body angle motions. New values are derived for the pitch semicircular canal acceleration threshold and response transmission time delay. The $0.05°/sec^2$ acceleration threshold is considerably less than those observed with subjective reports.

Linear motion sensors play no role in the detection of postural responses during free-fall divergence. Simulation of the otolith dynamic model indicates (19) that the linear acceleration threshold must be an order of magnitude less than the lowest values reported in the literature, even in the limit, as the initial body angle offset amplitude approaches zero. The utricle otolith organ threshold is sufficient to account for observed response threshold levels only in the static or nearly static range, since the body free-fall divergence rate (ω_B = 3 r/sec) is fast compared to the very slow dynamics of the otolith organs.

The role of linear and angular acceleration cues in posture regulation are made clear by observing the frequency ranges at which each component of the vestibular apparatus is most effective in responding to sinusoidal postural sway. Figure 8 shows the sinusoidal body angle amplitude necessary to achieve detection of motion as a function of sway frequency. Note the clear separation of the functional range for each sensor. The thresholds for each sensor within its frequency range of maximum sensitivity are approximately the same.

Postural Regulation with Vestibular Feedback

Posture control with vestibular feedback alone is observed. Subjects stand with eyes closed on the experimental platform. Ankle angle is maintained at zero by rotating the platform to track body angle motion. Ankle reaction torque and body angle are recorded continuously. A typical response sequence, is composed of the following stages:

1. stable period with no movement lasting several seconds
2. sway divergence begins
3. threshold is reached; ankle torque increases to restabilize body
4. one or several oscillations occur before quiet standing is re-established.

The transient response patterns are consistent among subjects. Average canal threshold is 0.36°. Variations among subjects are insignificant, $p > 0.1$.

The inverted pendulum configuration of the human body is stabilized with a combination of body sway and body angle and body angle rate feedback. Two modes of vestibular sensation provide dynamic feedback information of body angle. The semicircular canals provide a good estimate of sway rate for frequencies above 0.1 cps, while the otoliths indicate sway angle below 0.1 cps. Because of frequency response limitations of each sensor, the canal feedback control loop is unable to respond to rapid sway divergence. Thus, stability can

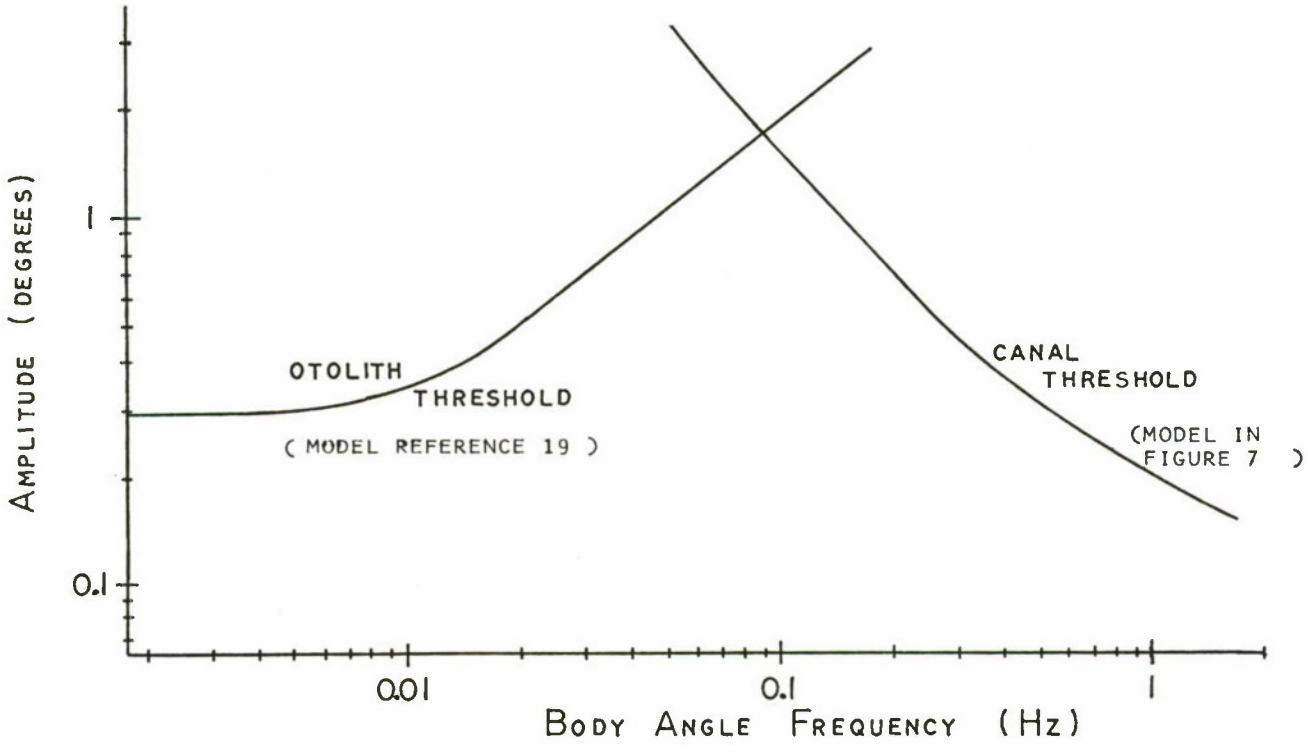


FIGURE 8 THE AMPLITUDE OF OSCILLATION JUST NECESSARY TO TO PRODUCE A THRESHOLD RESPONSE IN EACH VESTIBULAR ORGAN AS A FUNCTION OF SWAY FREQUENCY

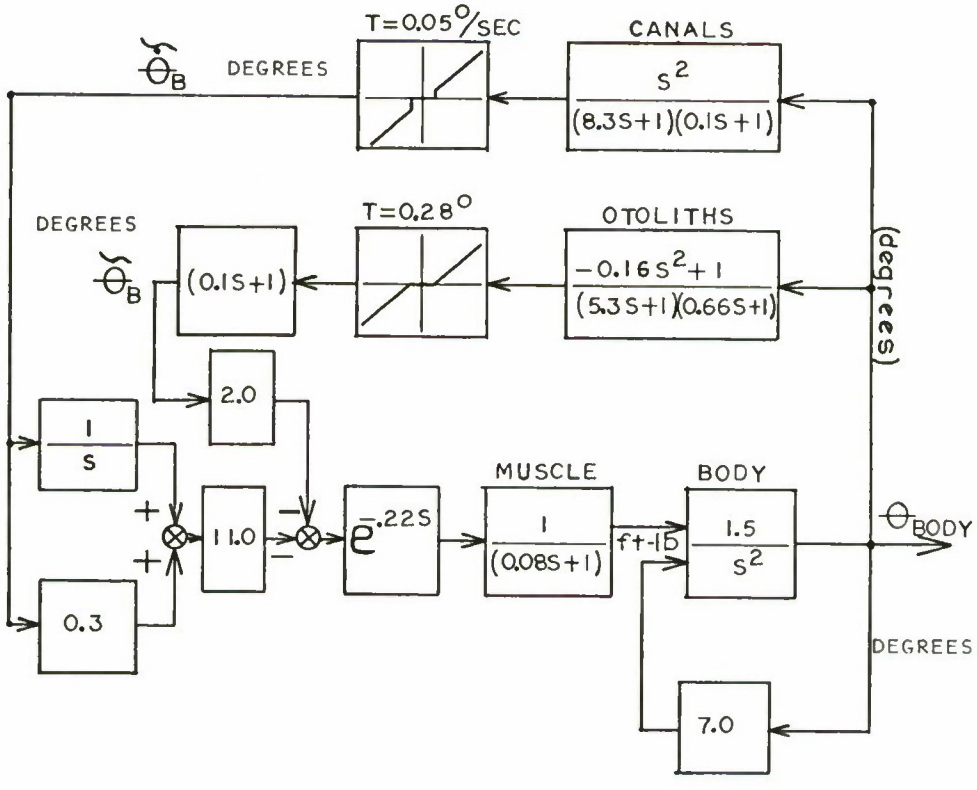


FIGURE 9 MODEL FOR VESTIBULAR REGULATION OF POSTURE

can only be attained through the combination of two frequency selective feedback loops, each of which in isolation is unable to provide stability.

In the proposed vestibular control model, the canal output, body angle rate, is used for an initial estimate of both body angle rate and, through neural integration, body angle. The low frequency utricle otolith estimate of body angle updates the initial canal estimate. The model configuration is shown in Figure 9.

The goal of the vestibular control model is to predict the response characteristics of a single cycle in the regulation process. Discrete response characteristics of the complete body-motor sensory model are shown in Figure 10. The presence of canal feedback alone shows the expected low frequency divergence. Stability is achieved when utricle otolith feedback is included to correct this. Note that the stimulation results can be compared to actual responses only during the first cycle, since no provision is made in the model to "reset" the vestibular thresholds.

4. Visual and Extereoceptive Senses

General models for visual and extereoceptive feedback are developed which indicate the ways that these senses modify vestibular responses. Analysis is limited to more general models for the following reasons:

1. In normal subjects visual and extereoceptive responses cannot be observed in isolation.
2. Present physiology does not permit construction of detailed models of the visual and extereoceptive senses.

Posture Regulation with Vestibular, Visual and Extereoceptive Senses

Effects of including visual and/or extereoceptive cues during vestibular control of posture are considered. Subjects stand on the experimental platform under the following conditions:

1. eyes open; platform rotates to maintain the ankle angle at zero degrees
2. eyes closed; platform rigid
3. eyes open; platform rigid.

In each of the above conditions, posture control follows the same basic pattern seen during vestibular feedback control: quiet period, body divergence, transient responses, and re-establishment of quiet stability.

The basic strategy of control in case 1, vestibular and visual cues only, is similar to control strategy with vestibular cues only. Control strategies in cases 2 and 3, however, are different, since the rigid platform enables full activation of the reflex control loop. Here control follows the pattern observed in the reflex response experiments - enhancement of reflex gain during transient disturbances. Subsequent analysis develops each of these control models to include effects of visual and extereoceptive cues.

Vestibular and Visual Control

Average threshold values for the vestibular response during divergence are slightly smaller when eyes are open, 0.29°, than when eyes are closed, 0.36°. In both cases, eyes open and eyes closed, variations among the three subjects are statistically insignificant, $p > 0.10$. Visual cues, however, do not appear to directly effect the vestibular response threshold. The body angle threshold for vestibular response is dependent on the divergence rate of the body, which is more rapid when eyes are closed. This effect is found to account for the difference between eyes open and eyes closed response thresholds.

Extereoceptive/Reflex Control

Basic changes in control are evident when extereoceptive cues and the reflex feedback loop are combined with vestibular regulation of posture. The following paragraphs explain these changes.

When reflex feedback control is included, the percentage of time during which subjects show static stability (i.e., when changes in reaction torque and body orientation cannot be detected) increases from an average of 20% to 40%. This increase is the result of "stiction" in the reflex loop.

The threshold for detection of body sway decreases significantly when extereoceptive cues are added, (case 2), dropping from 0.36° to 0.13°. Variations in the extereoceptive threshold value among subjects is statistically insignificant, $p > 0.10$. Amplitude of the transient response is reduced in proportion to the nearly threefold decrease in detection threshold.

Natural frequency of the transient response increases when the reflex feedback loop is added. Natural frequency with reflex stabilization is 0.35 Hz, significantly higher than that with only vestibular feedback control, o.20 Hz.

Addition of visual cues, (case 3), shows slightly improved damping characteristics during the slow phase correction. Threshold for detection of sway remains virtually unchanged.

Models for Extereoceptive and Visual Control

The model for reflex/extereoceptive control follows the strategy of gain control developed in Section 1. The model is shown in Figure 11.

The extereoceptive sense is modeled as a low threshold sway detector with habituation. Otolith cues provide slow phase correction due to habituation of extereoceptive cues. Because quiet standing with reflex and extereoceptive cues is inherently very stable, little effect is noted when visual cues are added.

The model is simulated. Figure 12 compares responses predicted by the model with experimental observations.

A visual feedback loop is added to the vestibular control model presented in Figure 9. The visual sense is modeled as a linear feedback controller with prediction (rate compensation) and transmission delay of 0.40 seconds.

The model assumes that visual feedback participates in posture control only intermittently. Visual control is suppressed during stable periods and activated only after vestibular detection of divergence.

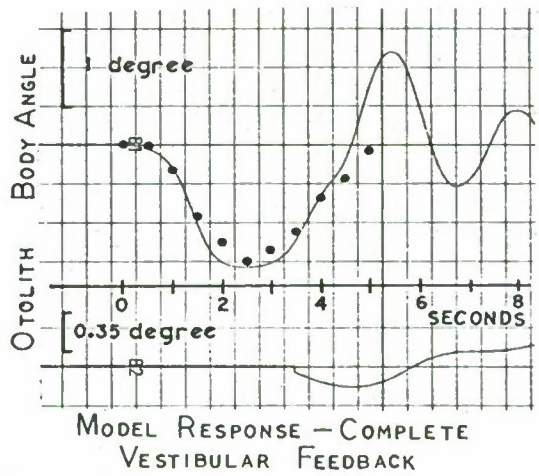


FIGURE 10 VESTIBULAR CONTROL MODEL CHARACTERISTICS COMPARED WITH EXPERIMENTAL DATA

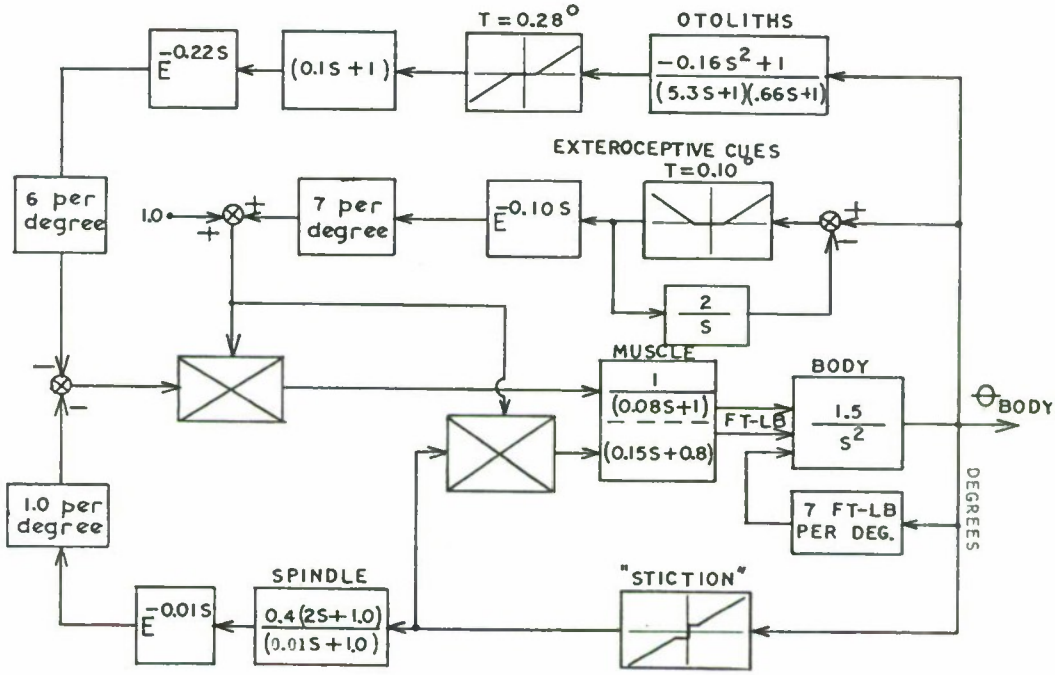


FIGURE 11 REVISED MODEL FOR REFLEX CONTROL OF POSTURE INCLUDING AND UTRICLE OTOLITH FEEDBACK LOOPS

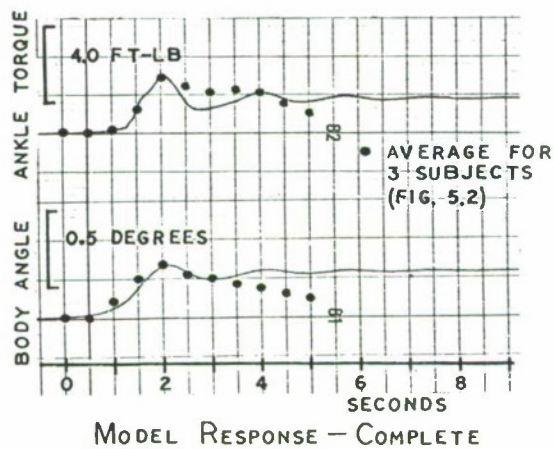


FIGURE 12 CHARACTERISTICS OF THE REFLEX/KINESTHETIC CONTROL MODEL COMPARED WITH EXPERIMENTAL DATA

The basic strategy of separating dynamic and static feedback control functions applies to extereoceptive and visual senses. Extereoceotive cues provide rapid, low threshold detection of body sway divergence. Visual cues effect well damped, high resolution correction of the slow phase drift.

5. Posture Control Without Vestibular Cues

Posture regulation of a subject with complete loss of vestibular function is observed. The subject, age twenty years, has complete loss of vestibular and auditory function due to bilateral transection of the eighth nerve. The subject's motor-sensory functions in the lower limbs are normal. Vestibular loss occurred about two years prior to the author's tests. Since loss of vestibulur function, the subject has remained active and has compensated for the sensory deficiency to the extent possible.

The Tests

The subject was observed standing quietly with eyes open and eyes closed. The following test was performed:

1. reflex response gains; eyes open and eyes closed
2. induced sway threshold tests; eyes open only
3. continuous recording of postural response and body angle motion.

Reflex Response Gains

Recalling earlier conclusions, the average gain of the stretch reflex response induced by small rotations of the ankle joint is shown to be about one-third that necessary for postural stability. During quiet standing with eyes open, the vestibular defective subject demonstrates reflex responses at gains somewhat larger than those of normal subjects, Figure 13. The level of gain, however, is still well below that necessary for postural stability. The subject is most likely using the same strategy for control as normal subjects.

Average reflex gain of the vestibular defective subject increases markedly when eyes are closed, Figure 14. The average reflex gain for extension, 12.35 ft-lb/deg, is large enough to achieve "rigid" postural stability.

This experiment tests the ability of the subject to detect sway divergence of reflex control and extereoceptive cues. Body sway is induced using backward and forward deflections of the platform at constant velocity. As sway is induced, the platform is rotated to track the rotation of the body, thus eliminating reflex and extereoceptive detection of sway via deflection of the ankle joints.

The observed response is similar to that seen in normal subjects: body angle begins to increase without a corresponding increase in ankle torque; motion is detected and torque increases rapidly; stability is re-established.

Threshold for Detection of Induced Sway

The minimum body angle subtended before motion is detected is about 0.35 degrees. The angle threshold increases continuously as the rate of induced sway increases, while normal subjects with the aid of vestibular cues show a constant threshold of 0.29 degrees for induced rates up to 0.80 degrees per second.

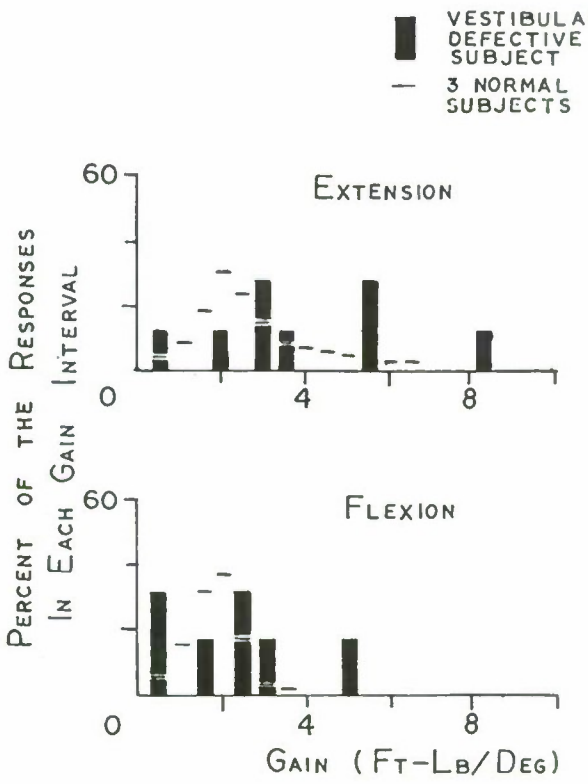


FIGURE 13 DISTRIBUTION OF REFLEX GAINS FOR 1/4 DEGREE STEPS, VESTIBULAR DEFECTIVE SUBJECT STANDING WITH EYES OPEN

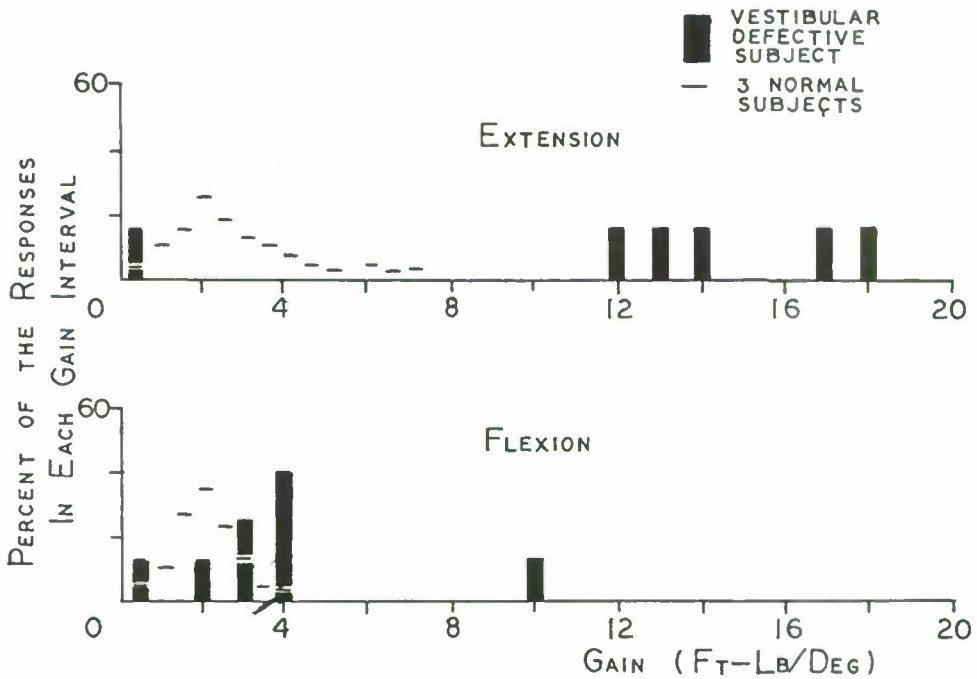


FIGURE 14 DISTRIBUTION OF REFLEX GAINS FOR 1/4 DEGREE STEPS, VESTIBULAR DEFECTIVE SUBJECT STANDING WITH EYES CLOSED

With eyes open, control strategy of the vestibular defective subject is the same as that observed for normal subjects. Records show periods of "stiction" stability and frequent diverging transients. Threshold for detection of the divergence transients is about the same as that observed for normal subjects, 0.10 degrees. Corrections of the transient disturbance, however are less consistent. Resulting divergence body angles are larger, 0.5 degrees versus 0.25 degrees for normals, and corrections are often underdamped.

When eyes are closed, the entire strategy of control changes significantly. No periods of "stiction" stability are present, rather small, higher frequency oscillations (about 3/4 to 1 Hz) are present almost continuously during quieter periods. During these quiet periods, body angle drifts continuously at rates ranging from 0.2 to 1.0 degrees per second. A rough estimate of the threshold for detection of this slow phase drift is about 2 to 4 degrees.

When slow phase drift is detected, larger transient responses are initiated. Characteristics of these responses are very erratic. Many of them are very poorly damped, showing many large oscillations at frequencies above 1 Hz.

Conclusions

A vestibular defective subject is able to regulate posture during a quiet standing, eyes open task using the same control strategy as normal subjects. A radical shift in strategy, however, is necessary to maintain stability when the vestibular defective subject closes his eyes. Results are consistent with the reflex/extereoceptive control model described in section 4. They confirm two major characteristics of the model:

1. Habituation of the extereoceptive gain control mechanism results in poor static stability of the reflex/extereoceptive control loop.
2. Gain of the reflex loop may be increased to achieve static "rigid" stability.

6. Conclusions

The posture control model can be subdivided into two basic parts: regulation with stretch reflex position feedback (standing on a rigid surface), and regulation relying on higher center feedback sensors (standing on a surface which rotates to track body angle motions, nulling ankle angle continuously.)

Posture control strategy in these two extreme cases is fundamentally different. Evidence is presented, however, which suggests that for a large class of conditions between these extremes of a perfectly rigid flat surface and a surface with special compliant properties, a combination of these two control strategies can be expected.

Posture Control on a Rigid, Flat Surface

During quiet standing on a rigid, flat surface, the ankle stretch reflex gains are about one-third that necessary for posture stability. Small "stiction" forces acting between fibers within both intra- and extra-fusal muscle, however, supplement this reflex gain, and together they provide a gain adequate for complete stability for very small ankle deflections.

Posture Control with Vestibular and Visual Senses

When reflex/extereoceptive feedback is removed, the subject must rely completely on higher center motion sensors, the vestibular and visual systems. With eyes closed, vestibular cues are sufficient to provide postural stability. In this case, the utricle otoliths and the semicircular canals operate as frequency selective feedback sensors. Canals, the higher frequency motion sensors, detect body divergence and initiate postural responses.

Posture Control without Vestibular Senses

Posture control without vestibular function is nearly normal when a defective subject stands on a rigid surface with eyes open. Detection of body divergence, extereoceptive cues, is normal. Reflex gains are somewhat higher than normal. It may be concluded that extereoceptive detection of divergence and visual correction of slow drift are sufficient.

When eyes are closed, a radical change in control strategy is evident. Since neither visual nor utricle otolith static senses are available, reflex gain is increased about sixfold to enable "rigid" stability.

An Overall Summary

Posture control is seen as a multiloop system in which a number of specialized feedback sensors contribute to the generation of commands. Proprioceptive sensors and neural processing at the lowest levels enable crude but fast acting responses based on information from body centered frames.

"Inertial" sensors and higher center processing provide more accurate, adaptable control but with longer processing delays. Hence, posture control is a highly non-stationary process in which responses to transient disturbances are initiated at the lowest levels. Allocation of control then "radiates" upwards to the higher centers where successive corrections based on more complete information, fine tune the initial responses.

Note: This paper is written as a brief review of Dr. Nashner's thesis.

References

1. Brown, M. C., Goodwin, G. M. and Matthews, P. B. C., "After Effects of Fusimotor Stimulation on the Response of Muscle Spindle Primary Afferent Endings," J. Physiol., Vol. 205, 1969, 677-94.

2. Clark, Brant, "Factors Influencing the Perception of Angular Acceleration in Man," Second Annual Status Report, NASA NGL-05-046-002, 1969.

3. Eccles, R. M. and Lundberg, A., "Supraspinal Control of Interneurons Mediating Spinal Reflexes," J. Physiol., Vol. 147, 1959.

4. Gernandt, B. E., Magdolna, I. and Livingston, R. B., "Vestibular Influences on Spinal Mechanisms," Experimental Neurology, Vol. 1, 1959.

5. Hill, A. V., "The Heat of Shortening and the Dynamic Constants of Muscles," Proceedings of the Royal Society of Britain," Vol. 126, 1938.

6. Hill, D. K., "Tension Due to Interaction Between the Sliding Filaments in Resting Striated Muscle. The Effect of Stimulation," J. Physiol., Vol. 199, 1968.

7. Houk, J. C., "Mathematical Model of the Stretch Reflex in Humans," S.M. Thesis, M.I.T., 1963.

8. Jansen, J. K. S. and Matthews, P. B. C., "The Central Control of the Dynamic Response of Muscle Spindle Receptors," J. Physiol., Vol. 161, 1962.

9. Johnson, A. R., "The Servo-Analysis of Postural Reflexes," Ph.D. Thesis, M.I.T., 1960.

10. Joyce, G. C. and Rack, P. M. H., "Isotonic Lengthening and Shortening Movements of Cat Soleus Muscle," J. Physiol., Vol. 204, 1969.

11. Joyce, G. C., Rack, P. M. H., and Westbury, D. R., "The Mechanical Properties of Cat Soleus Muscle During Controlled Lengthening and Shortening Movements," J. Physiol., Vol. 204, 1969.

12. Kim, J. H. and Partridge, L. D., "Observations on Types of Response to Combinations of Neck, Vestibular, and Muscle Stretch Signals," J. Neurophysiology, Vol. 32, 1969.

13. McRuer, D. T., Magdaleno, R. E. and Moor, G. P., "A Neuromuscular Activation Systems Model," Third Annual NASA_University Conference on Manual Control, 1967.

14. Matthews, P. B. C., "Evidence that the Secondary as Well as the Primary Endings of the Muscle Spindles May be Responsible for the Tonic Stretch Reflex of the Decerbrate Cat," J. Physiol., Vol. 204, 1969.

15. Meiry, J. L., "The Vestibular System and Human Dynamic Space Orientation," NASA-CR-628, 1966.

16. Partridge, L. D. and Kim, J. H., "Dynamic Characteristics of Response in a Vestibular Reflex," J. Neurophysiology, Vol. 32, 1969.

17. Roberts, T. D. M., Neurophysiology of Postural Mechanisms, Plenum Press, Butterworths, England, 1967, 90-120, 81-89, 120-175.

18. Webb, Paul, Bicastronautics Data Handbook, NASA SP-3006, 1964.

19. Young, L. R. and Meiry, J. L., "A Revised Dynamic Otolith Model," Aerospace Medicine, Vol. 39, 1968.

20. Nashner, Lewis M., "Sensory Feedback in Human Posture Control," Sc.D. Thesis, M.I.T., 1970.

EXPERIMENTAL AND GENERALIZED MATHEMATICAL ANALYSIS OF THE HUMAN MOTOR CONTROL SYSTEM

by Gerhard Vossius and Jürgen Werner

Physiological Institute, J.W.-Goethe-University, Frankfurt, W.-Germany

SUMMARY. Human motor activity is controlled by a highly organized system with manifold possibilities of adaptation, prediction and error-correction. The continuously operating system seems to be a typical extrapolation-system, which analyses the input-signals and their deterministic structure in the time-domain and generates a prediction for a time-interval adapted to the amount of relevant information. The discontinuously operating system can be approximated by a sampled-data-model, realizing that the sampling period may vary within a wide range and that the internal gain of the sampled-data-system can be altered by the system itself, following a strategy of optimization.

Human tracking movements have been analysed in special experimental setups using continuous and sampled input-signals. Regarding the results, we have to assume a typical predictor-system, which is able to compensate dead-time, to bridge signal-gaps or to continue interrupted signals. For the continuous part of movement we developed the following extrapolation-model:
The input f(t) is known up to the actual moment $t=t^*$ and can be approximated, if the derivatives exist, by a Taylor-series:

$$(1) \qquad \varphi(t, t^*, N) = \sum_{\nu=0}^{N} \frac{f^{(\nu)}(t^*)}{\nu!} (t-t^*)^{\nu}$$

By this we get a more or less exact extrapolation, depending on the signal-structure for $t = t^* + \Delta\tau$:

$$(2) \qquad \varphi(t^*+\Delta\tau, N) = \sum_{\nu=0}^{N} \frac{f^{(\nu)}(t^*)}{\nu!} \Delta\tau^{\nu}$$

But we should take into account, that a biological object like man will process all information disposible during a certain

interval Δt in order to carry out an extrapolation. This is done by the following model:

$$(3)\quad \Phi(t^{*}+\Delta\tau,\ \Delta t,\ N) = \int_{\Delta t}^{0} \sum_{\nu=0}^{N} \frac{f^{(\nu)}(t^{*}-\tau)}{\nu!}(\Delta\tau+\tau)^{\nu} g(\tau)\, d\tau$$

That means that extrapolations during an interval Δt are weighted by a function $g(\tau)$ and are summarized to one extrapolation-output. According to physiological facts, the weighting-function $g(\tau)$ may have the following form:

$$(4)\quad g(\tau) = \frac{e^{(\Delta t-\tau)/T} - 1}{-T\left(e^{\Delta t/T} - 1\right) + \Delta t}$$

The parameter T permits an individual adaptation of the function $g(\tau)$, whereas the expression in the denominator guarantees, that the following condition will be satisfied:

$$(5)\quad \int_{\Delta t}^{0} g(\tau)\, d\tau = 1$$

In another series of experiments we have shown, that the motor-control-system is able of identifying the deterministic structure of both continuous and discontinuous signals (which e.g. obey certain periodic laws) and by this to predict the signal in the course of the experiment by way of reproduction. Simple algorithms have been developed which simulate the identification-process using mechanisms of comparison of time-intervals and amplitudes.

A third very important principle obviously realized in the higher centers responsible for voluntary movements is discontinuous error-correction, the activity of which depends primarily on the momentaneous error.

The applicability of such theoretical aspects was examined by way of a simulation-program, which comprises properties of extrapolation, reproduction and error-correction. The model contains these three different branches, which cooperate in a ccordinated manner (cf. fig.1). The system, as a whole, is able to adapt the data-processing parts to different classes of input-signals. Fig.2 shows an example of eye-movement in experimental and computer recording.

In a special experimental setup for the eye-tracking-system, being the dominant control-loop for voluntary movement, some remarkable adaptive features of the discontinuous part of the system have been revealed.

We chose a step-function as target-signal and in addition we fed back the discontinuous (saccadic) eye-movement. The response was a sequence of steps. With an experimental feedback-gain $K_R \geqslant +1$ the system shows instability, whereas for negative feedbacks ($-10 < K_R < 0$) the system is able of avoiding instability by choosing an appropriate inner gain (fig.3). This is an unexpected effect and is in contrast e.g. to the sampled-data-model of Young and Stark. Defining the inner gain in the n-th sampling period $K_{vn} = \frac{\Delta a_n}{f_{n-1}}$ and calculating this gain for the sequence of steps, we have the typical procedure of fig.3:

A final value $K_{v\ opt}$ is reached after a few oscillations. In most cases K_{vn} does not oscillate about the final value, and generally a $K_{vn} \approx K_{v\ opt}$ is followed by an extremely inaccurate $K_{v\ n+1}$ larger than the final value. The error f_n, however, oscillates about its final value zero.

We interpret this, using terms of control-theory. Regarding fig.4 we find the following expression for the error

$$f_n = f_{n-1} \left\{ 1 - K_{vn} \, (1-K_R) \right\} \tag{6}$$

which enables us to determine the optimal gain $K_{v\ opt}$ for $f_n = 0$:

$$K_{v\ opt} = \frac{1}{1 - K_R} \tag{7}$$

As this is practically the inversion of the total feedback, it seems to be a relatively simple operation, which could be done in one step only. But, really, we should have in mind, that even without any external feedback K_R, the system, in most cases, will not react with the necessary gain K_{vn}. Now, introducing an additional feedback, this uncertainty of adjustment interferes with the process of identification of the external feedback and it is evident that this process will require a certain period.

As we stated already, the system prefers a changing sign for the error, being a plausible strategy, for only the final (desired) value $f_{n\ opt} = 0$ is known, whereas $K_{v\ opt}$ is unknown. Such a strategy, however, implies $K_{vn} > K_{v\ opt}$, regarding equation (6) and (7). If $K_{vn} < K_{vopt}$, we have

$$\text{sign } f_n = \text{sign } f_{n-1}$$

if, on the other hand, $K_{vn} > K_{v\ opt}$

$$\text{sign } f_n = - \text{sign } f_{n-1}$$

Using the observable variables Δa_n and f_n, $K_{v\ opt}$ might be identified in the following way:

$$(8) \quad K_{vopt} = \frac{-\Delta a_n}{f_{n-1}} = \frac{\Delta a_n}{f_{n-1}} \cdot \frac{1}{\left(1 - \frac{f_n}{f_{n-1}}\right)}$$

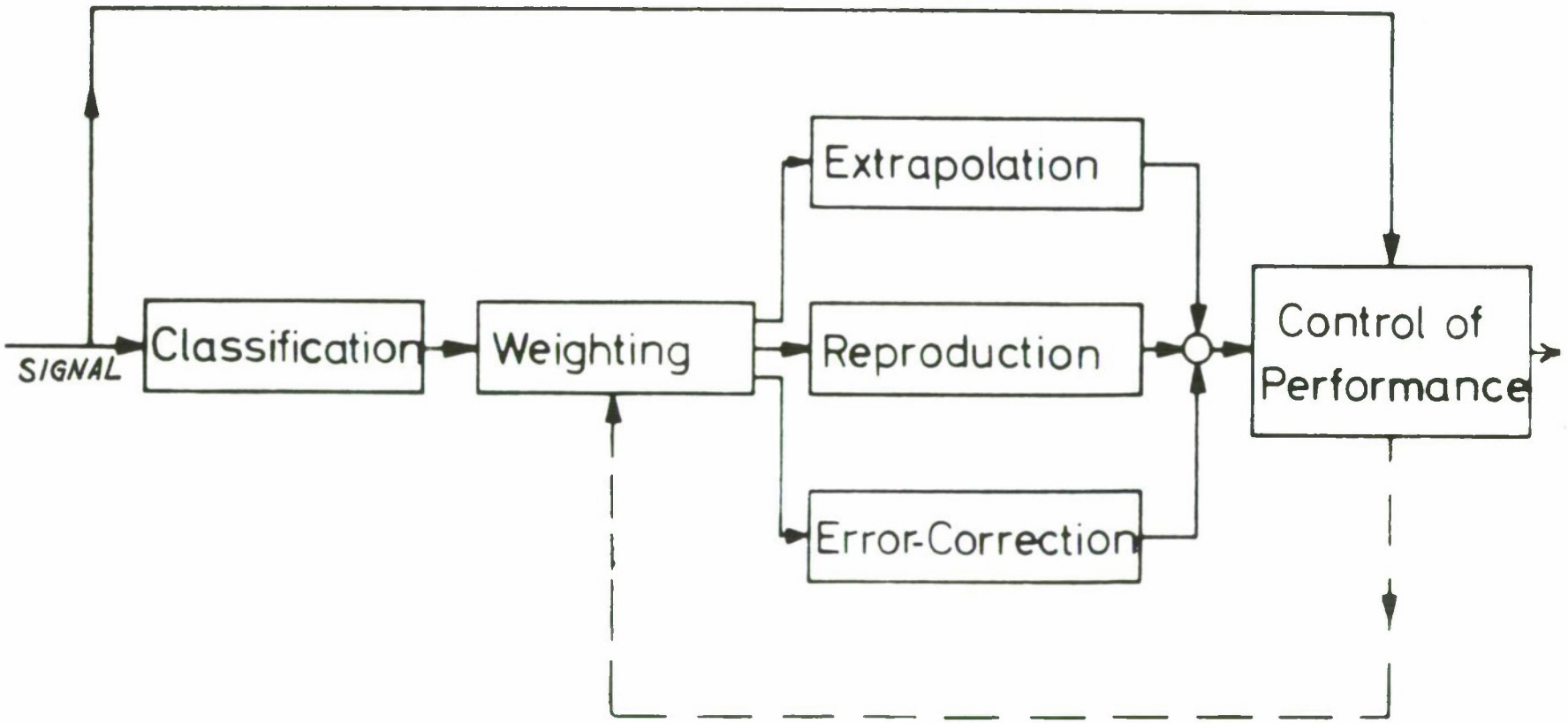
or with the "gains"

$$(9) \quad K_{vn} = \frac{\Delta a_n}{f_{n-1}} \qquad K_{fn} = \frac{f_n}{f_{n-1}}$$

$$(10) \quad K_{vopt} = \frac{K_{vn}}{1 - K_{fn}} \qquad -\infty < K_{fn} < 1$$

If now, without having identified $K_{v\ opt}$, Kvn had been close to $K_{v\ opt}$, we obtain $K_{fn} \approx 0$. Both K_{vn} and particularly K_{fn} are of small size, resulting in an inaccurate measurement and by this in difficulties of identification. On the other hand, $K_{vn} \rightarrow 1$ involves a big negative K_{fn}, so both K_{vn} and K_{fn} can be measured accurately, and a better identification of $K_{v\ opt}$ is achieved. Such a hypothesis is supported by the experimental fact, that the final value K_{vopt} generally is reached after a large deviation of K_{vn}.

Identification, adaptation and prediction are obviously very essential principles of motor control, and we do hope that with further development of theoretical research in system theory we shall have even more suitable instruments for analysing such complex biological systems.



Information-Processing of higher Centers

(Schematic Model)

Figure 1.

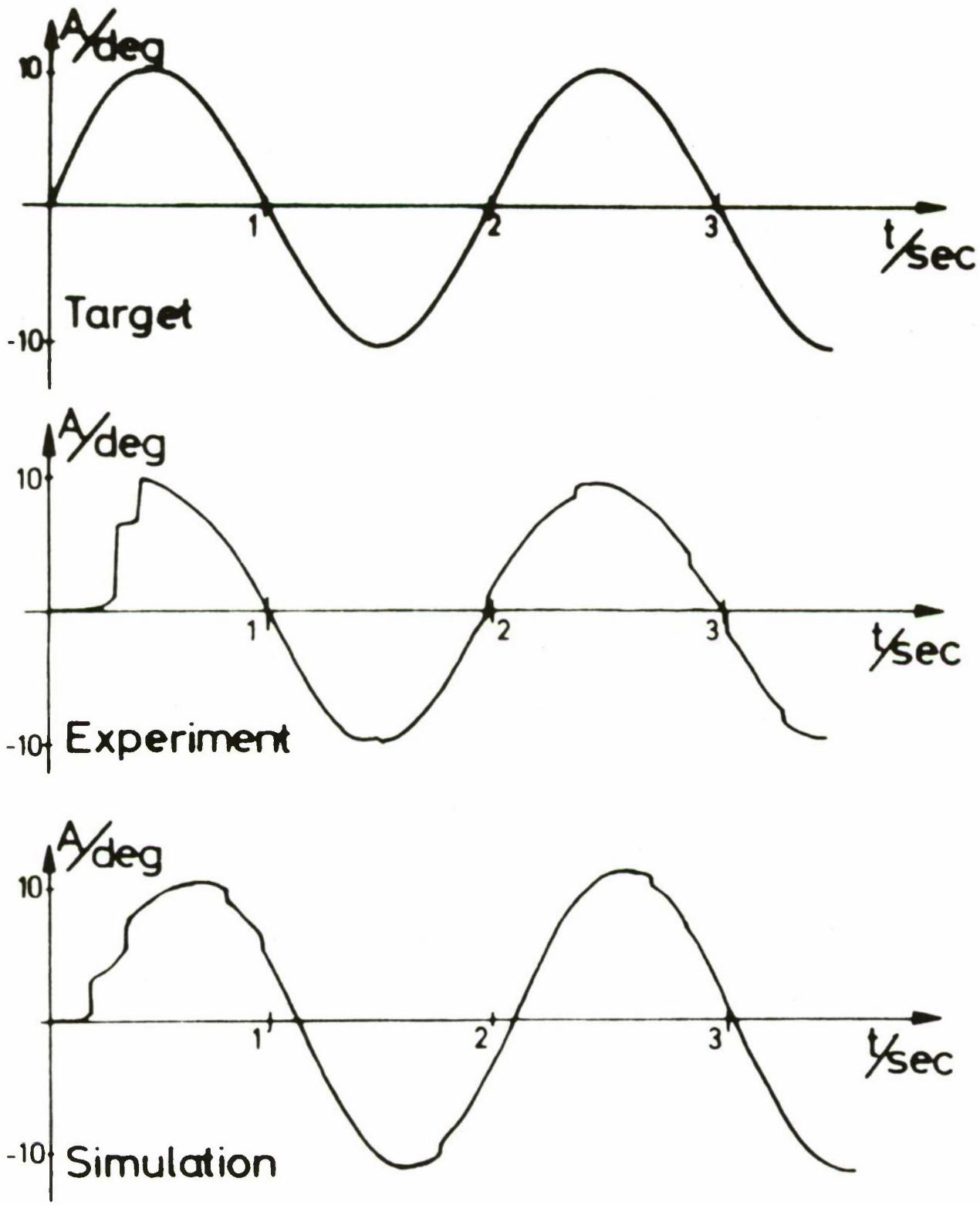
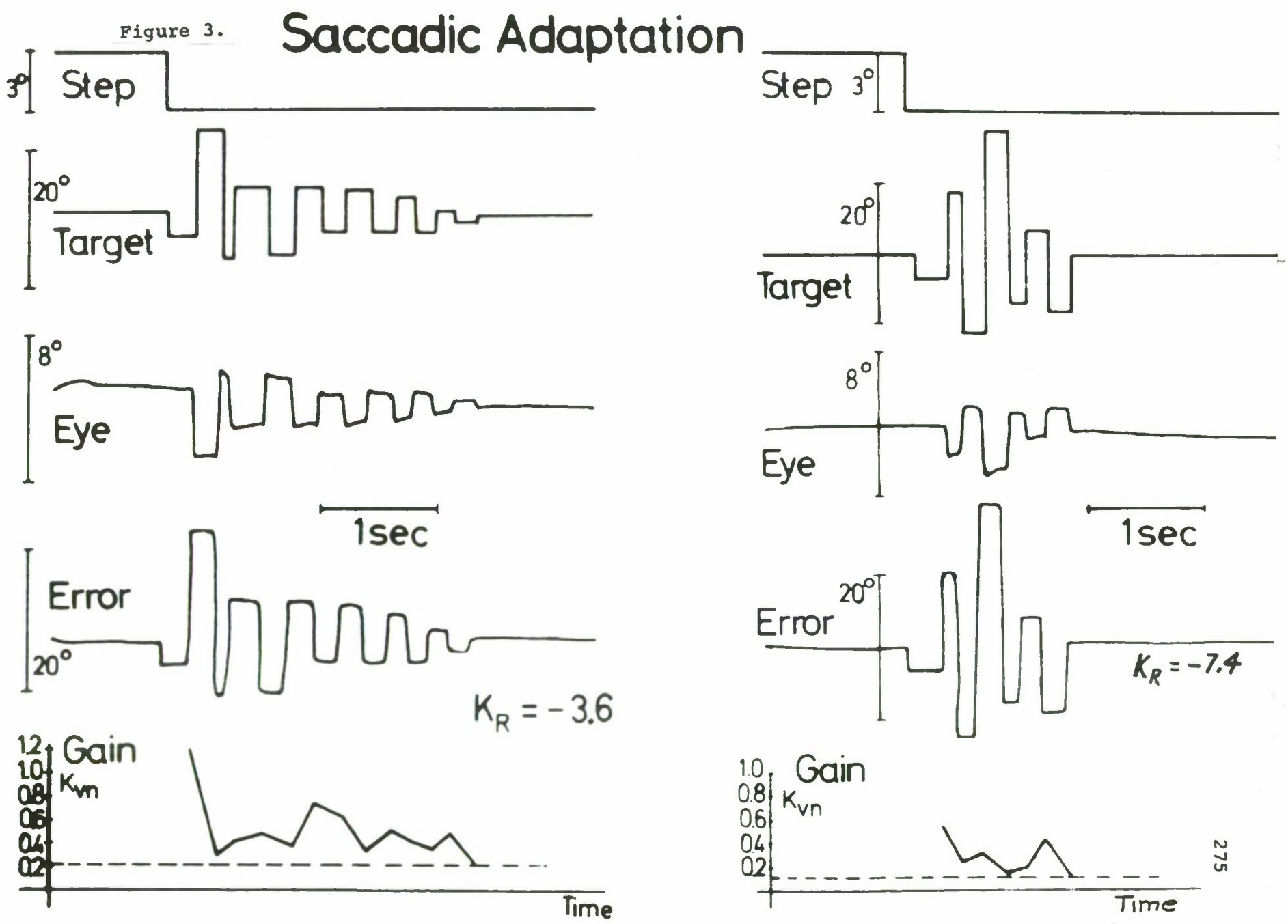

A/deg
10
-10
1
2
3
t/sec
Target
A/deg
10
-10
1
2
3
t/sec
Experiment
A/deg
10
-10
1
2
3
t/sec
Simulation

Figure 2.

Figure 3.


Saccadic Adaptation
Step
3°
20°
Target
8°
Eye
1sec
Error
20°
$K_R = -3.6$
Gain
K_{vn}
1.2
1.0
0.8
0.6
0.4
0.2
Time
Step 3°
20°
Target
8°
Eye
1sec
20°
Error
$K_R = -7.4$
Gain
K_{vn}
1.0
0.8
0.6
0.4
0.2
Time

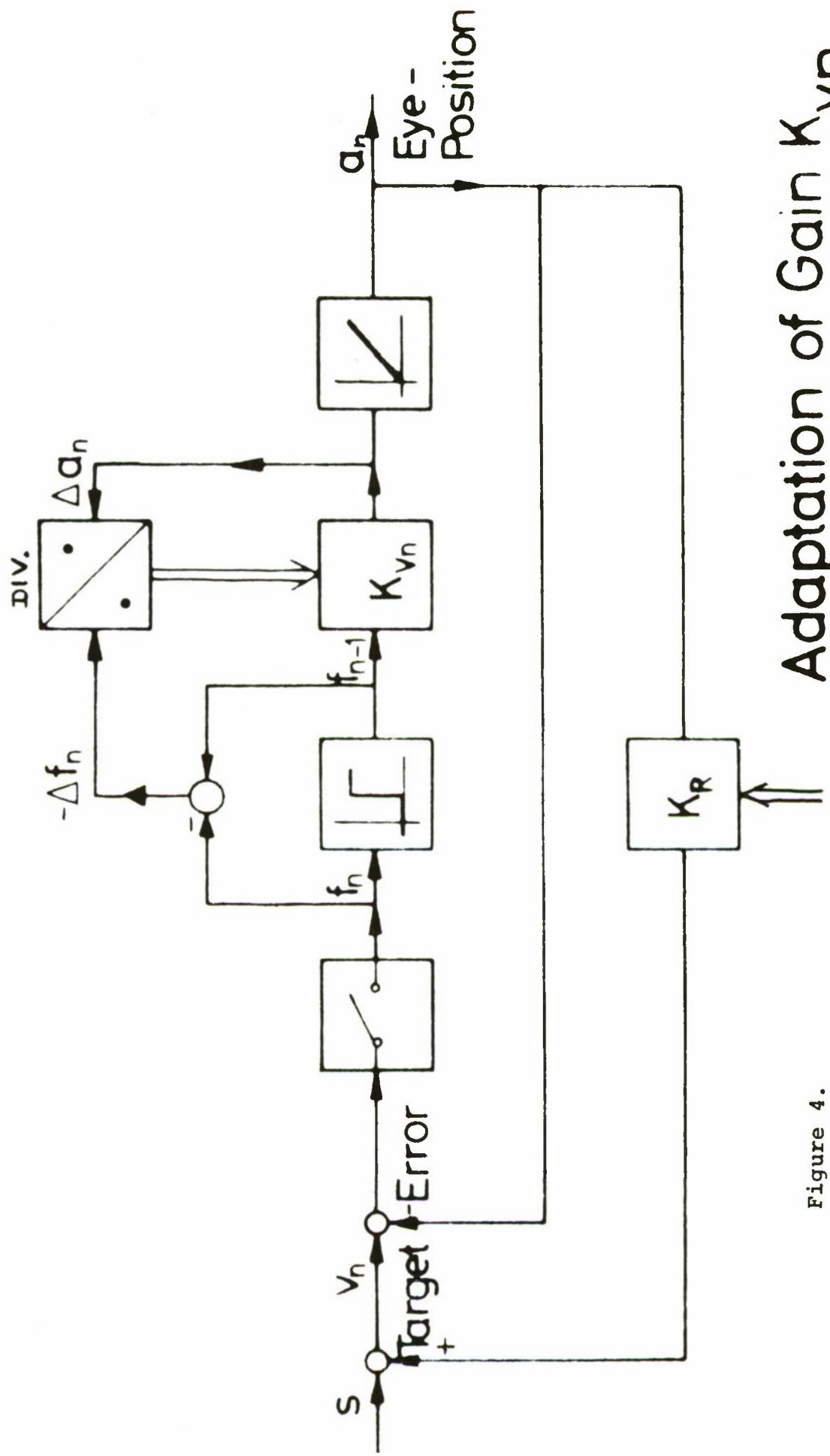

Adaptation of Gain K_{vn}
Eye-Position
a_n
Δa_n
DIV.
K_{vn}
f_{n-1}
$-\Delta f_n$
f_n
K_R
Error
Target
v_n
S

Figure 4.

DIRECT EVIDENCE AGAINST SAMPLING AT THE ALPHA MOTONEURON

Gyan C. Agarwal and Gerald L. Gottlieb
College of Engineering
University of Illinois at Chicago Circle, 60680
and
Biomedical Engineering Department
Presbyterian-St. Luke's Hospital
Chicago, Illinois 60612

ABSTRACT

Some results of a physiological investigation of the sampling hypothesis in normal human subjects are presented in this paper. The hypothesis of a proprioceptively open loop system at the initiation of voluntary effort is not supported by the data. No discontinuity in the monosynaptic pathway (primary afferent fibers from the spindles to alpha motoneuron) is observed in random step tracking under isometric conditions. This would indicate that the sampling hypothesis at the alpha motoneuron as proposed by Navas and Stark is not valid. The sampling behavior in the human motor system must be of central origin.

I. INTRODUCTION

The concept of the human operator performing tracking as a sampled data system has long been accepted. Craik [1] suggested that the response of the human operator in a control system is not continuous but intermittent. Following this work, Ward [2], Bekey [3], Lemay and Westcott [4], Lange [5], and others have postulated sampled data models for the human movement.

Stark and his collaborators [6-12] have made significant contributions towards the understanding of human motor system bridging the gap between engineering and physiological literature. In 1961, Stark [6] postulated variable topology of the neurological motor control system based on his "free-wheeling" experiments. The Navas and Stark model for pursuit tracking of unpredictable inputs is shown in Figure 1. The component enclosed in the dotted box is their alpha motoneuron (AMN) computer which replaces the sampler and hold of the other sampled data models. We will not go into their evidence for suggesting this model but some of their remarks are quite pertinent to the present paper and we quote:

"The behavior of the motor coordination system would occur as follows: Voluntary movements are preprogrammed in the higher centers and signaled as a whole. At the time of command release (position 1, Figure 1 of this paper) spindle afferent control is reduced and higher control of the alpha efferents is fully turned on. The system is operating open loop, characterized by high gain, low damping, and possible postural drift, during a short period of time, about 100 msec. Only apparent viscosity opposes motion; however, regulation comes from proprioceptive feedback and

from the apparent viscosity. After the motion has been substantially completed, the spindle afferent control of the alpha efferent pathway is fully turned on and the gamma efferent may be altered for a new set point; this corresponds to position 2 in Figure 1. The arm may oscillate for a time and is finally clamped to a new position. The new position remains stationary for approximately 230 msec." (*From Navas and Stark [11]*).

In none of the papers where sampled-data models of the human operator have been presented, has direct physiological evidence of the sampling mechanism in the motor system been shown. In Navas and Stark model, sampling or intermittency (as they have preferred to label it) at the AMN has been postulated based on indirect evidence. The Bekey and Elkind models place the sampler just after the summing point, and the Lemay-Westcott and Raoult models place it in the central processor after some lead [10].

In physiological literature no such sampling mechanism at the AMN has been shown in animals or in humans. As a matter of fact the physiological evidence seems to be against the sampling hypothesis at the AMN. If one accepts Merton's servo-control hypothesis of the motor system [13], it is inconceivable that the Ia-AMN synapse is disabled during voluntary control. Granit's concept of "α-γ linkage" control of the motor system [14] also seems to be contradictory to the disabling of the feedback path.

Because of the widely accepted theory of sampling behavior in human motor response, we chose to reexamine this issue in the hope of specifying the location of the sampler and the physiological mechanism involved. In a highly complex biological system such as the motor system, one can easily come up with a model to fit a limited set of data but the model may have little resemblance to the actual anatomical topology or to its physiological mechanism.

II. METHOD OF SYSTEM IDENTIFICATION

In this investigation, we do not wish to examine all of the proposed sampled-data models but will focus our attention to the Navas and Stark model for pursuit tracking of unpredictable inputs. Our aim is to examine if the postulated sampling behavior at the AMN does occur in human subjects. To identify the sampling behavior, it is sufficient to obtain some measure of the gain of the proprioceptive feedback loop (the inner loop in Navas and Stark model) as a function of time when a subject is executing a pursuit tracking task of an unpredictable target. Our identification procedure is similar to the one used by control engineers in the identification of a physical system in a continuous operating mode. Basically, one injects a short pulse input into a normally operating system and observes the transient behavior. The input pulse should be short enough and of low intensity so as not to disturb the on-going process very significantly. Although this process of identification is not ideal, it is at least something physically possible.

The neurophysiologist uses a somewhat similar approach for identification of certain neural activities of the spinal cord. A schematic diagram of the neural connections for the proprioceptive feedback loop indicating only the extensor muscle around the ankle joint is shown in Figure 2. If one considers the movements around this joint, Figure 2 represents part of the neural pathways associated with the gastrocnemius-soleus muscles (GSM). The identification of a monosynaptic connection between the largest afferent fibers (Ia) and spinal motoneurons [15,16] gave to the neurophysiologists a new important tool for the experimental investigation of the reflex activity of the spinal cord. Observations of Hoffmann [17] concerning the reflex response of the motoneurons by electrical stimulation of the mixed tibial nerve in the popliteal fossa in human subjects is a hallmark of human neurophysiological investigations. Later investigators, Magladery, Paillard, and others [18-24], have greatly developed this H-reflex technique (so named after Hoffmann).

The explanation of the Hoffmann reflex and associated activities is as follows: 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 GSM. About 32 msec after the stimulus, a synchronized electromyogram (EMG) burst, the H-wave, is recorded from the GSM and simultaneously with the H-wave, the isometric tension begins to rise. The H-reflex has been shown to be monosynaptic [19]. As the level of stimulus is increased, the H-wave 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 the 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. Typical responses to progressively stronger stimuli are shown in Figure 3. The explanation for this behavior is that the Ia afferent nerves conduct the stimulus volley centripetally, faster than the antidromic volley is conducted in the alpha motor fibers. Thus, the H-wave is blocked by a collision of the antidromic wave going up with the reflex wave as it travels orthodromically down the alpha motor fibers.

If one uses a stimulus of submaximal strength (H-wave with no M-wave in the response), the peak-to-peak amplitude of the H-wave provides an estimate of the reflex excitability of the monosynaptic arc. For a constant input stimulus level, this reflex gain may be measured at different points in time when a subject is executing a repeatable voluntary task. This identification scheme provides a means for testing the sampling hypothesis at the alpha motoneuron.

III. EXPERIMENTAL PROCEDURE

The experiments were performed on five human subjects ranging in ages from twenty-one to thirty-six years. A subject was seated normally in a chair with his right leg extended, the knee slightly flexed and the foot strapped to a fixed plate having attached four strain gauges in a bridge circuit for measuring torque. The signal from the strain gauge bridge circuit was used to provide a visual display of the level of foot torque (FT) on one of the channels of a dual beam scope. The position of a target (second spot on the CRT) was controlled by a voltage set by an IBM 1800 computer that ran the experiments and recorded the data. A schematic of the apparatus is shown in Figure 4. The computer sampled and recorded four output data channels, a) foot torque at 250 SPS (samples per second); b) full wave rectified and filtered (third-order Paynter filters [25]) electromyographic activities (EMG) recorded using surface electrodes on the two muscle groups, gastrocnemius-soleus (GSM) and anterior tibial (ATM), at a rate of 500 SPS; and c) unfiltered EMG from the GSM for the H-wave at 1000 SPS for 50 msec after the stimulus. Electrical stimuli were applied to the subjects tibial nerve by means of cutaneous electrodes located posteriorly in the popliteal fossa and anteriorly above the knee. The electrodes were held in place by straps. The stimuli were monophasic, square pulses of 1.5 msec duration provided at 7-10 second intervals by a Grass S-8 stimulator with an SIU5 isolation unit. All electrodes were coated with *Sanborn Redux Creme.*

The basic idea of the experiment is to test the excitability of the Ia-alpha motoneuron synapse over the full time course of a repeated voluntary effort by means of an H-reflex elicited at different times during such efforts. The subject is instructed to track a target which jumps back and forth at random 7-10 second intervals between two positions. The subject tracks the target with his own spot by changing the amount of force he is exerting on the foot plate. He is told to make his movements rapid but not to anticipate the jump of the target. After the target jumps to a new position, the stimulator is triggered by the computer with a delay specified by the experimenter. By varying this delay and allowing for the natural variability of the subject's response time, the H-reflex could be elicited over the entire interval of the voluntary movement.

IV. RESULTS

At submaximal stimulus inputs, the peak-to-peak amplitude of the H-wave has been shown to be a function of the static muscle activity i.e., the static or initial foot torque (IFT) (Agarwal and Gottlieb-manuscript in preparation). The H-wave amplitudes recorded, when the subject is tracking a step input by changing his muscle activities, have to be corrected for this static influence on H-wave to record dynamic facilitation or inhibition of the motoneuron pool. Figures 5 and 6 show

the results of two typical experiments. For each experiment there are eight figures coded "A" through "H", which are as follows:

A. This shows the variation in H-wave with static or initial foot torque (IFT). The solid line in the figure shows the piecewise linear normalizing curve fitted to the data by eye.

B. This shows to what degree the data of part A deviates from the normalizing curve. The equation for H-wave variation (ordinate in B) is

$$\Delta H = \frac{H_m - H_n}{H_n}$$

where H_m is the measured value of the H-wave and H_n is the value of the normalizing curve at the same value of torque.

C. This shows the amplitude of the H-wave measured during step tracking. In this and the succeeding parts of these figures, the symbol "+" will denote a measurement made during a dorsiflexional step effort and the symbol "x" will denote a measurement made during a plantarflexional step effort. The solid line shows the normalizing curve of part A.

D. The time at which the H-reflex is recorded is under computer control but the time of the initiation of the step effort varies with each data record. Therefore, a variable TAI (Time After Initiation) is defined which is the interval of time between the initiation of the effort and the recording of the H-wave. Part D shows the value of the foot torque (FT) in step tracking at the times when H-wave was recorded.

E. The amplitude of the filtered voluntary EMG cannot be measured at the same instant at which the FT in part D is recorded because the H-wave and the stimulus artifact are much larger than the voluntary activity signal and are measured on the same pair of wires (see Figure 7). Therefore, the voluntary EMG is sampled 40 msec before the H-wave, about 8 msec prior to the stimulus artifact. Part E shows the voluntary activity (averaged over 20 msec) of the anterior tibial muscle during dorsiflexion versus TAI. Note that the time axis has been shifted by 40 msec.

F. This shows the voluntary activity (averaged over 20 msec) of the soleus muscle during plantarflexion versus

TAI. Again, note that the time axis has been shifted by 40 msec.

G. This shows the actual value of the H-wave plotted against TAI.

H. This shows the H-wave variation ΔH during dynamic tracking (part G) with respect to the normalizing curve of part A. The positive and negative values of the ratio ΔH denote facilitation and inhibition, respectively, due to dynamic tracking, with respect to the static value of the H-wave for the same foot torque.

The data GCA073 is from a subject of average height and light build, and the data HBN007 is from a subject of average height and stocky build. These figures show the typical variations observed during some 50 experiments of this nature. Figure 7 shows two responses recorded directly from a CRT display (which only the experimenter could see). Marked facilitation during plantarflexion and inhibition during dorsiflexion is clearly shown. Note also that the on-going step effort is not significantly disturbed by the pulse input.

V. DISCUSSION

Before discussing the results of our experiments, some differences between the present experimental procedure and that of Navas and Stark should be made clear. Whereas Navas and Stark's work was done on the motor control system used for wrist rotation which involves pronator and supinator muscles, we have investigated the movements around the ankle joint which involve gastrocnemius-soleus and anterior tibial muscles. The physiological roles for these systems are different and so are their mechanical characteristics. However, the reflex mechanisms are apparently the same. Most of Navas and Stark's work was done with no external loading on the movement, while our experiments were done under isometric conditions of the joint. The ankle joint was chosen because of the accessibility of the tibial nerve which allows for testing the alpha motoneuron pool by a monosynaptic reflex mechanism. Also, the EMGs from the muscles can be easily recorded via surface electrodes. The isometric condition was chosen to minimize any artifact due to electrode displacement. In all the experiments, part A of Figures 5 and 6 was repeated after the tracking experiment. Whenever any significant differences in curves before and after tracking experiment beyond the normal variance in H-wave were observed, that experiment was discarded.

The influence of static foot torque on the amplitude of H-wave (Figures 5A and 6A) has been discussed in a separate paper (manuscript in preparation). Plantarflexion facilitates the motoneuron pool so that more motoneurons are brought synchronously above threshold by the Ia volley. Conversely, dorsiflexion inhibits the motoneuron pool. The dynamic facilitation and inhibition is a very potent effect as shown in

Figures 5C and 6C. Note in Figure 5D the step effort is nearly complete in about 300 msec from the time of initiation and is symmetrical in plantar and dorsal directions, whereas, in Figure 6D, the plantarflexion effort is complete in about 160 msec and dorsal effort takes about 240 msec.

In Figure 5H, the gastrocnemius-soleus (GS) motoneuron pool has dynamic facilitation all during plantar movement. The facilitation is shown at the initiation of the effort and decreases as the effort is completed. During the dorsal dynamic effort, the GS motoneuron pool is inhibited throughout and maximum inhibition is at the time when the ATM is most active.

In Figure 6, since the plantar effort is completed in about one-half the time used by the subject GCA, the dynamic facilitation of the motoneuron pool is very strong initially, followed by some dynamic inhibition and then facilitation again. The inhibition of the GS motoneuron pool for a short interval may be due to the Golgi tendon feedback because the GSM is contracting very vigorously about that time. During the dorsal movements, the GS motoneuron pool shows continuous dynamic inhibition. Similar results were obtained in other experiments.

These experiments indicate that the monosynaptic pathway (Ia to AMN) of the GS motoneuron pool is continuously active during the plantarflexion effort in unpredictable (time of occurrence) step effort tracking. Dynamic facilitation at the initiation of the plantar effort is very potent (figure 6H). This contradicts the Navas and Stark hypothesis that at the initiation of the voluntary movement, the system is operating proprioceptively open loop. Certain implications of the present findings concerning the mechanisms of voluntary control have been discussed in a separate paper [26].

Also note that no discontinuity in the monosynaptic pathway is observed during the course of the plantar movement which would indicate that the sampling does not occur at the AMN or any other peripheral point. Consequently, the so-called sampling behavior in the human motor system must be of central origin. However, at this stage, it is not clear whether the sampling hypothesis is necessary to explain the limited amount of data (peak in the gain characteristics in complex sine wave tracking, step-like response in slow ramp tracking, and short width unpredictable pulse response) which supports sampling behavior. Some nonlinearity in the system may give rise to these characteristics. Much more has to be learned about the human motor system before this question can be settled.

Acknowledgment

This work was partially supported by NIH Training Grant and PSL Grant-in-aid.

VI. REFERENCES

1. K. J. W. Craik, "Theory of the Human Operator in Control Systems," Brit. J. Psychol., Vol. 38, pp. 56-61, 1947; pp. 142-148, 1948.

2. J. R. Ward, "The Dynamics of a Human Operator in a Control System: A Study Based on the Hypothesis of Intermittency," Ph. D. Dissertation, Dept. of Aeronautical Engrg., University of Sidney, Austrailia, 1958. (Cited by Bekey).

3. G. A. Bekey, "The Human Operator as a Sampled-data System," IRE Trans. on Human Factors in Electronics, Vol. 3, pp. 43-51, 1962.

4. L. P. Lemay and J. H. Westcott, "The Simulation of Human Operator Tracking Using an Intermittent Model," International Congr. Human Factors in Electronics, Long Beach, California, 1962.

5. G. W. Lange, "Representation of the Human Operator as a Sampled-data System," Proc. IEE, London, Vol. 115, pp. 342-354, 1968.

6. L. Stark, "Neurological Organization of the Control System for Movement," QPR-61, Res. Lab. Elec., MIT, pp. 234-238, 1961.

7. L. Stark, Y. Okabe, and P. A. Willis, "Dynamic Characteristics of Motor Coordination System in Man," Biophysical J., Vol. 1, pp. 279-300, 1961.

8. J. C. Houk, "A Mathematical Model of the Stretch Reflex in Human Muscle Systems," M. S. Thesis, MIT, 1963.

9. F. Navas, "Sampling or Quantization in the Human Tracking System," M. S. Thesis, MIT, 1963.

10. L. R. Young and L. Stark, "Biological Control Systems - A Critical Review and Evaluation - Developments in Manual Control," NASA CR-190, 1965.

11. F. Navas and L. Stark, "Sampling or Intermittency in Hand Control System Dynamics," Biophysical J., Vol. 8, pp. 252-302, 1968.

12. L. Stark, Neurological Control Systems - Studies in Bioengineering, Section 5: The Hand, Plenum Press, N. Y., 1968.

13. P. A. Merton, "Speculation on the Servo-control of Movement," in The Spinal Cord, Ciba Foundation Symposium; Little, Brown and Company, Boston, pp. 247-260, 1953.

14. R. Granit, Receptors and Sensory Perception, Yale University Press, New Haven, 1955.

15. B. Renshaw, "Activity in the Simplest Spinal Reflex Pathways," J. Neurophysiol., Vol. 3, pp. 373-387, 1940.

16. D. P. C. Lloyd, "Conduction and Synaptic Transmission of the Reflex Response to Stretch in Spinal Cats," J. Neurophysiol., Vol. 6, pp. 317-326, 1943.

17. P. Hoffmann, Untersuchungen uber die Eigenreflexe (Sehnreflexe) Menschlichen Muskein, Jul. Springer Verlag, Berlin, 1922.

18. J. W. Magladery and D. B. McDouglal, 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 Fibres," Bull. Johns Hopkins Hosp., Vol. 86, pp. 265-290, 1950.

19. J. W. Magladery, W. E. Porter, A. M. Park, and R. D. Teasdall, "Electrophysiological Studies of Nerve and Reflex Activity in Normal Man. IV. The Two-neurone Reflex and Identification of Certain Action Potentials from Spinal Roots and Cord," Bull. Johns Hopkins Hosp., Vol. 86, pp. 499-519, 1951.

20. J. W. Magladery, "Some Observations on Spinal Reflexes in Man," Pflügers Archiv, Vol. 261, pp. 302-321, 1955.

21. J. Paillard, Reflexes et Regulations d'origine Proprioceptive Chez l'Homme, Librairie Arnette, Paris, 1955.

22. J. Paillard, "Functional Organization of Afferent Innervation of Muscle Studies in Man by Monosynaptic Testing," Amer. J. Physical Medicine, Vol. 38, pp. 239-247, 1959.

23. R. F. Mark, J. M. Coquery and J. Paillard, "Autogenetic Reflex Effects of Slow or Steady Stretch of the Calf Muscles in Man," Experimental Brain Res., Vol. 6, pp. 130-145, 1968.

24. H. Taborikova and D. S. Sax, "Motoneurone Pool and the H-reflex," J. Neurol. Neurosurg. Psychiat., Vol. 31, pp. 354-361, 1968.

25. G. L. Gottlieb and G. C. Agarwal, "Filtering of Electromyographic Signals," Amer. J. Physical Med., Vol. 49, 1970.

26. G. L. Gottlieb and G. C. Agarwal, "Interactions Between the Voluntary and Postural Mechanisms of the Human Motor System" (to appear in Journal of Neurophysiology).

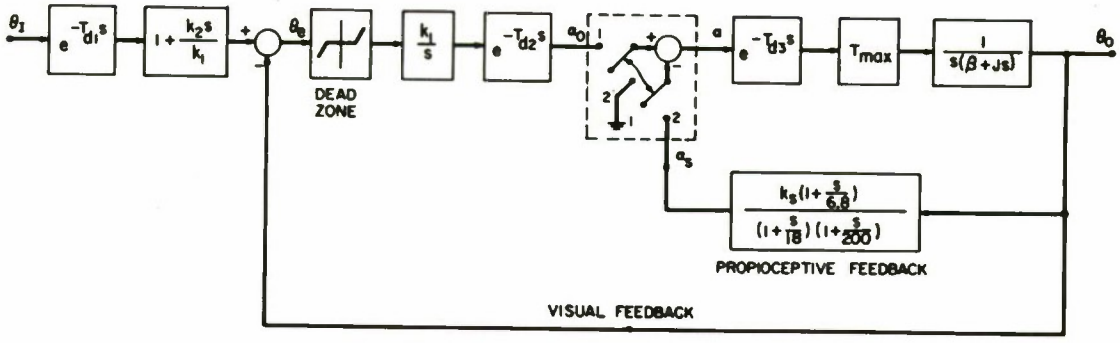


Figure 1. Navas and Stark model for pursuit tracking of unpredictable inputs.

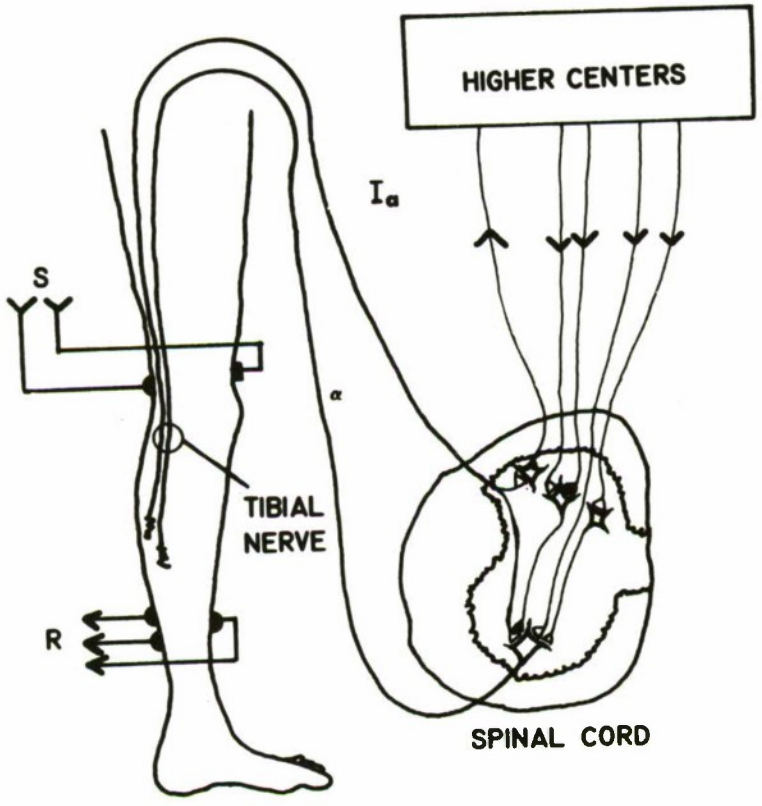


Figure 2. Neural pathways for H-reflex and supraspinal control. Stimulation is applied unifocally by S to the posterior tibial nerve in the popliteal fossa, and the reflex pathway runs up to the spinal cord by Ia afferent fibers and out to the gastrocnemius-soleus muscle by α efferent motor fibers. Other fibers of the tibial nerve are not shown. Recording R is from the soleus muscle.

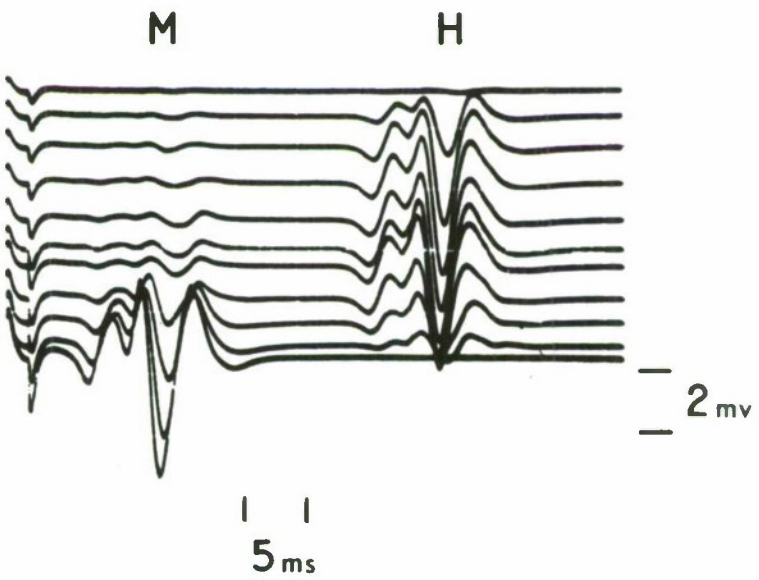


Figure 3. Typical responses to progressively stronger stimuli applied to the posterior tibial nerve (Subject GLG). At the far left is the stimulus artifact. The H-wave appears about 32 msec and the M-wave about 8 msec after stimulation. The stimulus is increasing from top to bottom.

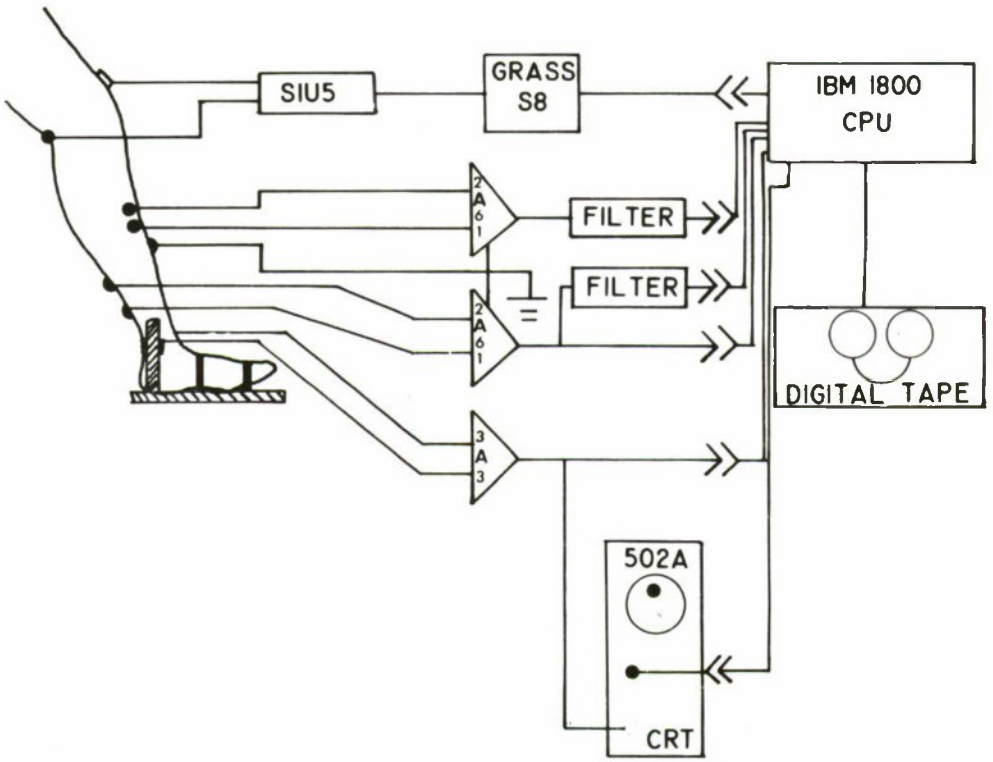


Figure 4. Schematic diagram of the apparatus. The components are Grass S8-stimulator, SIU5-isolation unit, 2A61-high gain a.c. differential amplifier (tektronix), 3A3-high gain d.c. differential amplifier (tektronix), FILTER-EMG full wave rectifier with third order Paynter filters, 502A-dual beam scope (textronix) for pursuit tracking, and IBM 1800 process computer.

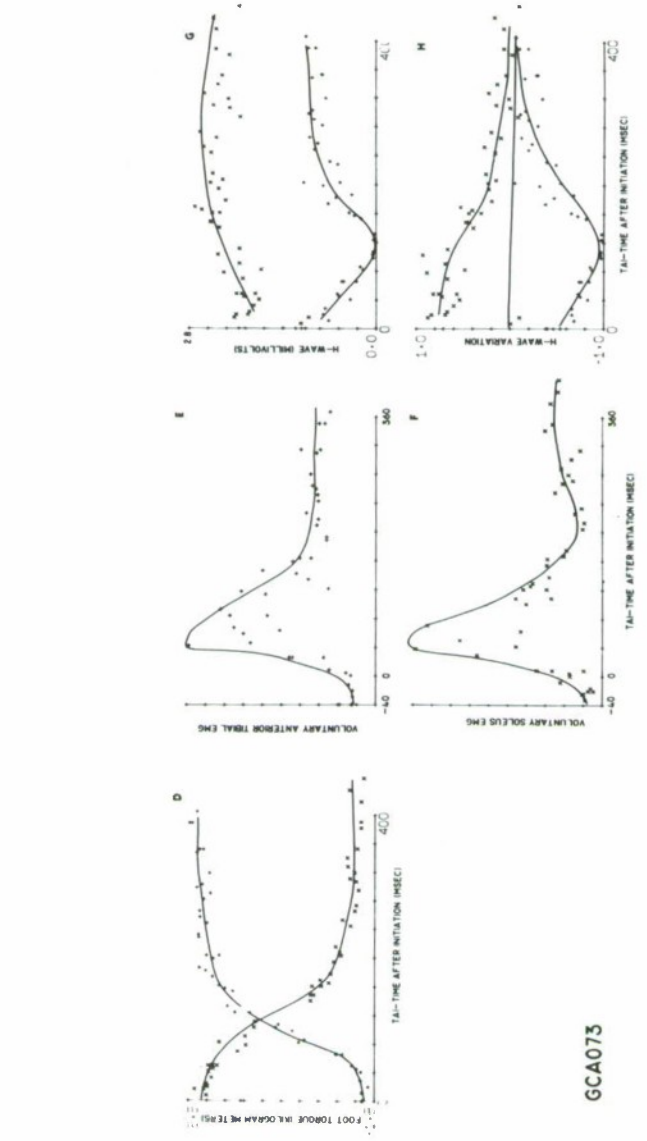
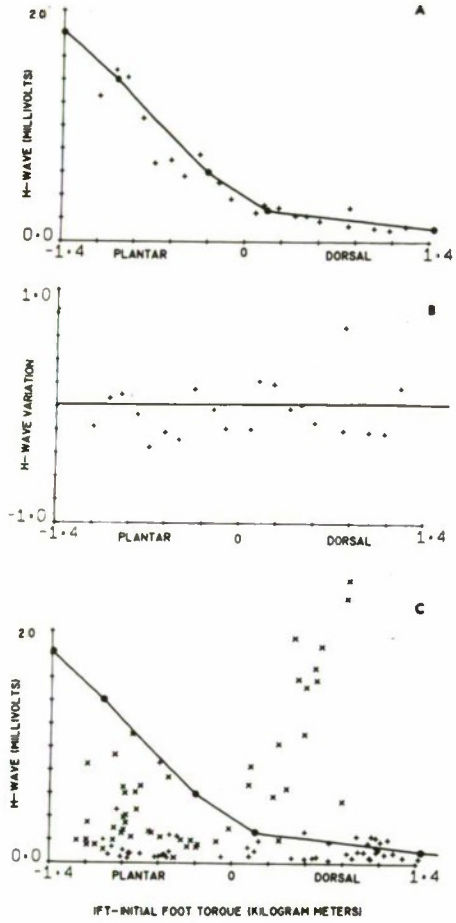


Figure 5. Response of subject GCA in unpredictable (time of occurrence), isometric effort step tracking (see text for explanation). Moderate input stimulus.



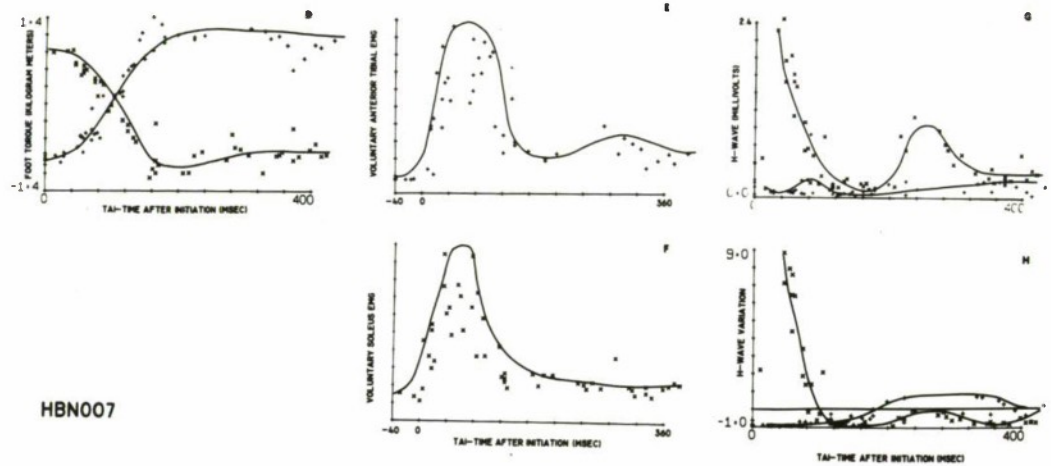


Figure 6. Response of subject HBN in unpredictable (time of occurence), isometric effort step tracking (see text for explanation). Low input stimulus.

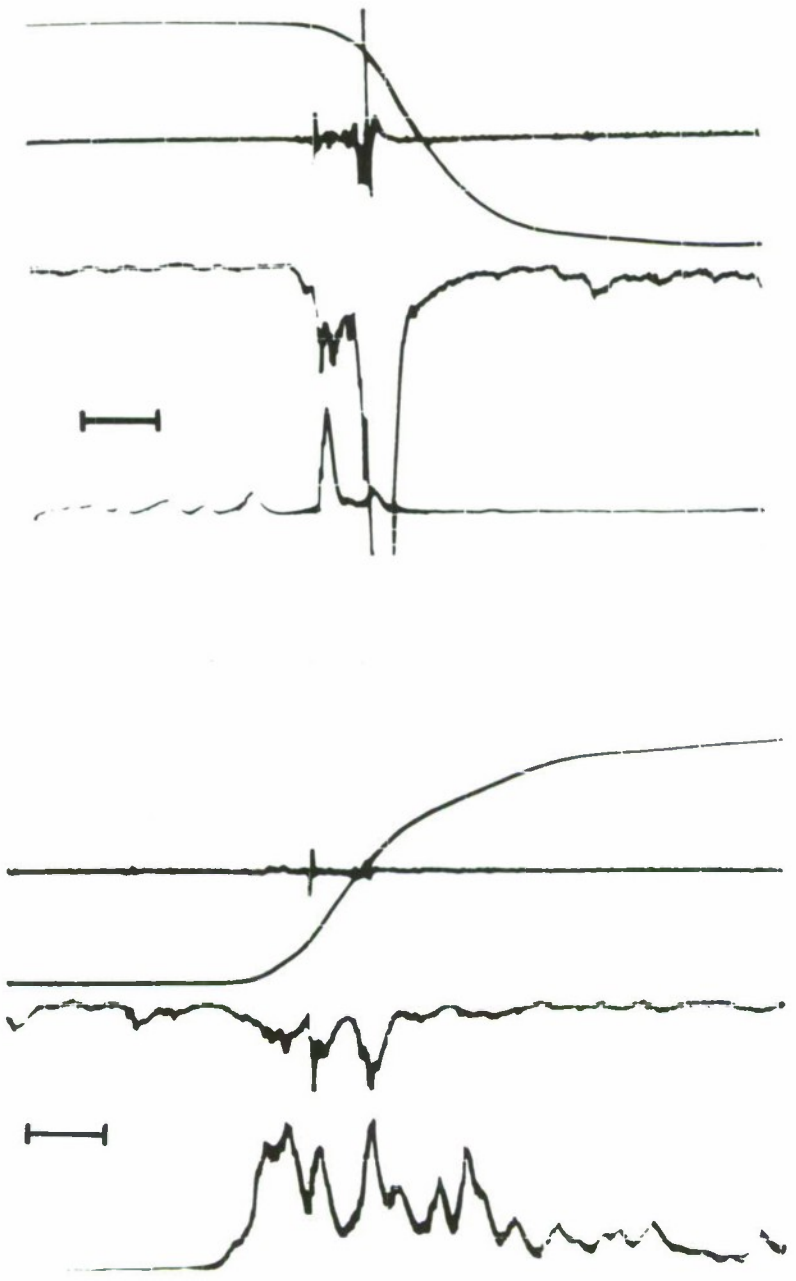


Figure 7. Typical individual responses in isometric step tracking. The four traces from top to bottom are: foot torque (1.2 kg.m/unit); H-reflex (0.8 mv/unit); processed EMG of GSM; and processed EMG of ATM (2.0 v/unit). The downward foot torque deflection is in the plantar direction. Strong facilitation during plantar effort and nearly total inhibition during dorsal effort are clearly shown. The first marker in the H-reflex trace is stimulus artifact which is also reflected in both processed EMGs. Time scale is 50 msec/unit. (Subject GCA)

APPLICATION OF PARAMETER OPTIMIZATION TO STABILITY AUGMENTATION DESIGN*

R. L. Stapleford
Systems Technology, Inc.
Hawthorne, California

ABSTRACT

A systematic procedure for the design of aircraft stability augmentation systems is proposed. The key features of this procedure are the selection of essential feedbacks from an examination of several handling quality metrics and the use of parameter optimization techniques to determine the numerical values of the SAS parameters. The optimization problem is structured to include both manual (pilot describing function model) and SAS feedbacks. The cost function includes pilot tracking errors and SAS control deflections. A method of selecting the relative weighting is presented.

The feasibility of this procedure is demonstrated by applying it to the longitudinal axis of the F-4 aircraft. Three widely different flight conditions were selected. For all three, the same SAS form (pitch rate and normal acceleration feedbacks to the elevator), the identical problem formulation, and the same method of selecting the cost function weights are used. The resulting systems are judged quite satisfactory and easily meet the short-period requirements of the current military handling qualities specification.

INTRODUCTION

Background

The material presented here is the result of a research program which had as its major objective the development of a more systematic procedure for designing stability augmentation systems. To provide such a procedure a design criterion which insures good handling qualities must first be specified. After the design criterion has been established, an optimization operation can be performed to determine a SAS mechanization which is compatible with that criterion. To establish the bases for a specific systematic procedure major portions of the research program were devoted to the criterion question and to study, assessment, and comparison of competing optimization techniques. The design process which evolved after decisions were taken on these points was tested in some sample applications for a longitudinal SAS for the F-4 aircraft. The results for these rather simple example applications were very encouraging. We therefore feel that the design process proposed here has considerable potential.

---

*This research was accomplished under Contract F33615-69-C-1359 for the Air Force Flight Dynamics Laboratory.

The design criterion proposed here is to minimize the pilot tracking errors for an outer loop command input. These errors are minimized for a pilot describing function model which has fixed lag elements and a variable gain. The cost function includes control deflection due to the SAS, in addition to the tracking error. The relative cost function weight is based on the aircraft control response at a selected frequency.

Given the design criterion there is still the question of the type of optimization procedure to use. To resolve this question, two different optimization techniques were considered, linear optimal control theory and parameter optimization. Our conclusion was that with the current state of the art, parameter optimization methods are superior to linear optimal control theory for use in a SAS design procedure. Since this conclusion is a key point in the proposed design procedure, and subject to dispute, a fairly detailed comparison of the two optimization techniques was presented in Ref. 1, Appendix A. Some of the major points of that discussion are summarized below.

A major problem in applying linear optimal control theory to flight control is the conversion of the optimal control into a practical optimal (or near optimal) controller. The optimal controller can be realized in a variety of ways ranging from pure gain feedback of all the state variables to all control points to a few feedbacks with rather high order equalizations. These results must be converted to practical controller mechanizations in which a relatively few, easily sensed quantities are fed back through fairly simple equalizations. Because the optimal controls are usually computed for several combinations of flight condition and vehicle configurations, another practicalization difficulty is to provide a means of adjusting the controller parameters so as to approximate the optimal solutions over the entire vehicle flight envelope. Both of these problems are major obstacles to the utilization of linear optimal control theory in a SAS design procedure.

On the other hand both of these problems can be avoided with parameter optimization techniques. In parameter optimization the controller form is specified, and then the computer determines the optimal values for these parameters. Since the controller form is prespecified only practical system mechanizations are considered from the outset. Only easily sensed parameters are used as feedbacks, only reasonable equalization schemes are used, and limitations on the parameters can also be imposed. These limitations might be set to avoid such practical mechanization problems as excessive structural feedback or oversensitivity to sensor noise. The problem of scheduling the controller parameters can also be avoided since the parameter optimization can be carried out simultaneously for a variety of flight conditions and vehicle configurations. In fact, a potential scheduling form can be specified and then the optimization program can determine the optimal values of the scheduling parameters.

Parameter optimization also offers several other advantages relative to linear optimal control theory. These include:

1. The final controller configuration is truly optimized in the sense that the optimal values of all available parameters are determined.

2. The inclusion of actuator and sensor dynamics is straightforward and presents no problems.

3. Constraints on certain control parameters such as feedback gain or lead-lag ratio can be specified.

4. Among the parameters which can be optimized are accelerometer location and rate gyro orientation.

5. Plant and controller nonlinearities can be included.

6. Nonquadratic cost functions can easily be utilized.

Proposed Design Procedure

As noted above, the proposed design procedure is based on pilot tracking error and SAS control deflection in the performance criterion and parameter optimization techniques. This procedure can be divided into three steps: (1) selection of essential feedbacks, (2) single case optimization, (3) multicase optimization. Each of these steps is outlined below.

To select the essential feedbacks, the basic handling qualitites deficiencies of the aircraft must first be identified. This is done by examining several key handling quality metrics for a wide variety of flight conditions. Once the deficiencies have been identified, the selection of practical feedback alternatives to improve the characteristics is rather straightforward. At this point the designer should have one or more practical and economically efficient controller configurations which are inherently capable of providing good vehicle handling qualities. Reference 1 describes this step in considerable detail.

Having selected potential controller forms, the next step is to perform a variety of single case optimizations. By "single case" we mean one combination of flight condition (Mach number, altitude, and normal acceleration) and vehicle configuration (weight and c.g. location). This step has two objectives. The first is to determine whether additional system complexity is required or whether the system can be simplified. The second is to make a preliminary estimate of the requirements for controller parameter scheduling. For those parameters which appear to require scheduling a preliminary estimate of the form of the scheduling algorithm should be made.

The final step is the multicase optimization. In this step the controller is optimized for a combination of flight conditions and vehicle configurations. The cost function would be the weighted sum of the cost functions for each individual case. The parameters to be optimized would include controller parameters which are held constant over the flight envelope, parameters in the scheduling algorithm, and sensor location and/or orientation. These optimizations can best be done on a large scale hybrid computer with the vehicle dynamics simulated on the analog portion, in fast time, and the optimization logic controlled by the digital portion of the computer.

The multicase optimization would be an iterative process. For each iteration, the acceptability of the system optimized for the range of

conditions would be assessed with regard to the individual flight conditions. Whether the system is satisfactory or not, the optimization would be repeated. If the system is satisfactory, attempts to develop a simpler system would be made. If it were unsatisfactory, additional complexity, such as additional feedbacks or more complicated scheduling, would be introduced. This iterative procedure would terminate when the designer became fairly confident of having achieved a minimum complexity, satisfactory system.

The remainder of this paper discusses the selection of the design criterion and the example application.

DESIGN CRITERION

General Problem Structure

In general, a controller optimization problem can have the following elements (see Fig. 1):

- Plant and controller
- Inputs
- Model
- Error filters
- Cost function

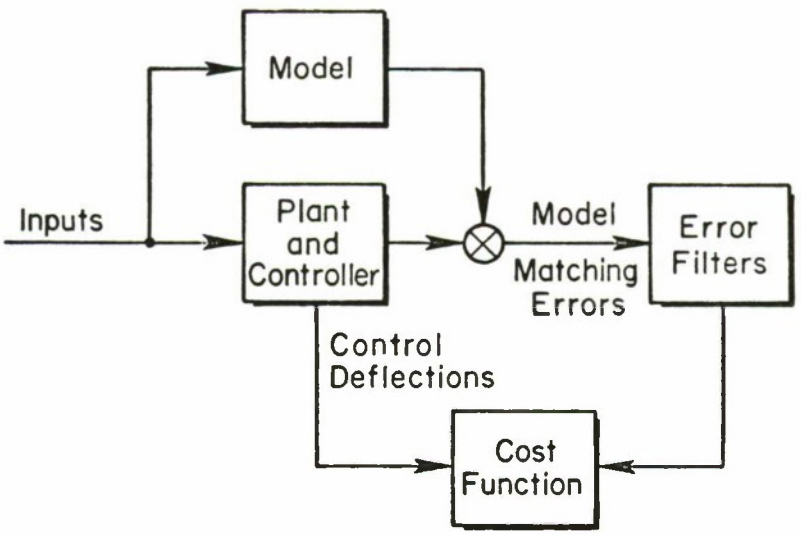


Figure 1. General Problem Structure

The physical elements in this list are the inputs, which represent the control environment, and the plant and controller, which constitute the control system. The remaining elements, model, error filters, and cost function, together amount to a system specification; usually they are not physically in the final system realization. Our present objective is to specialize this general structure to one appropriate for a SAS design procedure. To gain insight in what this might be, we will start by considering the effects of changes in each of the various elements. These, hopefully, will point the way to a specific SAS design structure.

Plant and Controller

A key problem in the system description is the definition of the plant and controller. This includes, as a minimum, the aircraft and the SAS but may or may not incorporate the pilot and associated manual control feedbacks. To distinguish between the two possibilities hereafter, the term PIL (pilot in the loop) will be used when manual feedbacks are included and POOL (pilot out of the loop) when manual feedbacks are not present. The relative advantages and disadvantages of these two plant and controller structures are summarized in Table I and discussed below.

TABLE I

RELATIVE MERITS OF PIL AND POOL STRUCTURES

PLANT AND CONTROLLER STRUCTURE	ADVANTAGES	DISADVANTAGES
PIL (Pilot in the loop)	1. Can use simple models 2. More directly compatible with basic pilot requirements	1. Complicates multicase optimization program 2. SAS/aircraft combination can be unstable
POOL (Pilot out of the loop)	1. Fewer parameters to optimize 2. More directly compatible with current handling qualities specifications	1. Optimal model-matching system may be unacceptable 2. Can be overly conservative with regard to SAS/aircraft stability

One advantage of PIL is that simple model forms can be used in the specification elements. These can be as simple as unity gain for command inputs and zero gain for disturbance inputs. The simplicity is due to the direct inclusion of the manual feedbacks in the problem. On the other hand, with POOL we need a model of the SAS/aircraft dynamics, which if adequately matched, would permit satisfactory manual control. In multiloop tasks (e.g., path outer loop and attitude inner loop) appropriate models for each loop are required.

The other advantage of PIL is that it is more directly compatible with basic pilot requirements. These requirements are adequate command following and disturbance suppression without excessive pilot lead. The restriction on pilot lead can be included in the optimization program by limiting pilot lead to those values which have little effect on pilot ratings (roughly 1 sec).

The major disadvantage of PIL is that it complicates the multicase optimization program. The complication arises because different kinds of restrictions are imposed on SAS and pilot parameter variations between cases. Restrictions on the SAS parameters will be one of the following:

- Constant over all cases
- Varied according to a completely specified algorithm
- Varied according to an algorithm of specified form (numerical values of which are to be optimized)

By contrast, there can be no such restrictions on the pilot parameters. Instead, they should be optimized for each flight condition according to the adjustment rules of pilot/vehicle system analysis. These are basic behavioral characteristics of the pilot which, in a sense, satisfy his internally specified dynamic performance criteria.

There are two solutions to the practical problems of multicase optimization with PIL. One is to include in the list of parameters to be optimized as many sets of pilot variables as there are different cases to be considered. This approach, while relatively straightforward, can significantly increase the number of parameters to be optimized. The second solution is to use a two-step optimization program which operates as follows:

1. Initial estimates of pilot and SAS parameters are inserted as inputs for each case.
2. The program does a multicase optimization of the SAS parameters while holding the pilot parameters fixed.
3. The program optimizes the pilot parameters for each case separately.
4. Steps 2 and 3 are iterated until the overall optimal solution is obtained.

The other PIL disadvantage is only a potential problem and has not arisen in any examples to date. Since the PIL structure imposes no direct restraints on SAS/aircraft dynamics, theoretically the SAS/aircraft combination could be unstable, with the pilot closures providing a stable overall system. This problem can probably be avoided by putting certain restrictions on the SAS parameters.

Let us now consider the advantages of POOL. First, there are fewer parameters to be optimized as the pilot parameters are omitted. Second,

this problem structure is more directly compatible with existing handling qualities specifications. All current specifications consider only SAS/aircraft dynamics and not pilot/SAS/aircraft characteristics.

The major problem with POOL is that the controller evolved from an optimal model-matching solution is frequently unacceptable. We tried this problem structure for the longitudinal augmentation of the F-4 at two different flight conditions. The results for both were completely unacceptable, see Ref. 1. With the relatively simple models which were used, the optimal solutions always had impractical SAS characteristics, e.g., extremely high gains or the wrong sign on the acceleration gain. To avoid this problem, one must use more complex models which represent fairly accurately the dynamics of the aircraft and a good augmentation system. To derive such a model for each flight condition is roughly equivalent to designing the SAS, so why bother with an optimization procedure?

The second disadvantage of POOL is the requirement that the SAS/aircraft system be stable. This is inherently necessary if the cost function integration is carried out over a long time interval. Thus, acceptably slow divergent spiral modes would not be tolerated. However, this problem can easily be circumvented. A fixed, low-gain feedback could be added to stabilize the spiral mode. The design procedure could then be applied to the augmented aircraft. Finally, the initial feedback could be deleted from the system.

After consideration of the above factors, it was decided that the PIL problem structure had considerably more promise than POOL for use in a practical SAS design procedure. The difficulties involved in specifying an adequate model for the POOL structure make that approach unattractive. Consequently, most of our effort was devoted to using the PIL form and it is the one recommended in our proposed design procedure. The remainder of this paper will consider only this form.

Cost Function

A quadratic type cost function is:

$$J = \sum_i \int_0^\infty [q_i x_i(t)]^2 \, dt \qquad (1)$$

where x_i is any pilot/SAS/aircraft response variable

q_i is the cost function weight on x_i

This is a particularly convenient form. With a linear constant coefficient system the cost function can be evaluated without having to integrate the equations of motion (see Ref. 1, Appendix B). Other types of cost function can be used with parameter optimization and, if the optimization is done on a hybrid computer, a wide variety of cost functions could be applied with little difficulty. However, for most problems the quadratic form is adequate and computationally convenient, especially for digital optimization. For these reasons, only the quadratic form of Eq. 1 will be considered further.

Given the cost function form, one still has the problem of picking which response variables to include and their relative weights (the q_i's in Eq. 1). Clearly, the model-matching errors should be included, but are additional terms necessary? Several examples were run with only an error term in the cost function. In some cases this resulted in extremely large feedback gains. With a pitch damper, large gain makes the pitch/elevator response look like K/s out to high frequencies. Consequently, nice pilot closures are achieved and the error response at low frequencies is extremely small. However, the damping of the high frequency modes in these instances was obviously too low, e.g., damping ratio of 0.02 and frequency of 38 rad/sec.

A simple, straightforward method of restricting the feedback gains is to include control deflection in the cost function. When using the PIL problem structure, this approach would restrict both pilot and SAS gains, whereas we are only interested in restricting the SAS. Pilot gains are adequately limited by closed-loop stability because of fixed lag elements in the pilot describing function. Consequently, instead of using total control deflection in the cost function we used only the control deflection due to the SAS feedbacks.

With two terms in the cost function there is the problem of selecting the relative weights. For guidance in solving this problem we borrowed an idea from linear optimal control theory, Ref. 2. The basic idea is as follows:

- Given a model reference optimization where the plant output, x, is to follow the model for frequencies up to ω_1
- The relative cost function weights on control deflection, δ, and model-matching error, should be

$$\frac{q_\delta}{q_\epsilon} \doteq \left| \frac{x}{\delta} (j\omega_1) \right| \tag{2}$$

This concept was employed to the problem at hand and the frequency, ω_1, was selected as approximately that of pilot/vehicle system crossover. Example applications indicated that a frequency of 4 rad/sec was generally satisfactory.

For longitudinal SAS design with manual response to a pitch command, the relative weights used were*

$$\frac{q_{\delta_{SAS}}}{q_\epsilon} \doteq \left| \frac{\theta}{\delta_e} (j4) \right| \tag{3}$$

*For simplicity, the magnitude of the $|\theta/\delta_e(j4)|$ asymptote was used rather than the actual value.

For some situations, it would be more appropriate to consider manual closure of path loops, as well as attitude inner loops, and path commands. Then it would probably be better to use a frequency typical of manual crossover for a path loop, roughly 0.5 – 1 rad/sec.

Input Characteristics

For all the examples which were run, the input command was assumed to be of the form

$$\theta_c = \frac{1}{s + p_i} \tag{4}$$

In selecting the value of p_i, Eq. 4 should be interpreted as the transient analog of a random input with power spectra

$$\Phi_{\theta_c\theta_c}(\omega) = \frac{k}{\omega^2 + p_i^2} \tag{5}$$

rather than as a deterministic input. Viewed in this light, p_i should be set at the bandwidth of typical pitch commands. About 1 rad/sec is an appropriate order of magnitude for this. Much lower values were also considered for a number of cases, but the results were judged inferior. Therefore a value of 1 rad/sec was chosen.

For situations containing manual path control, it would probably be desirable to reduce the input frequency. Path command and disturbances are generally considerably lower in bandwidth than their attitude control equivalents. For these cases a value of roughly 0.5 rad/sec is suggested as a starting point.

Model Dynamics

For command inputs the simplest possible model is unity gain. Since this proved to be satisfactory in the example applications, it was selected. However, brief investigations were also conducted with two other models. One was

$$\left(\frac{\theta}{\theta_c}\right)_{model} = \frac{3}{s + 3} \tag{6}$$

This is the closed-loop transfer function for

$$Y_p Y_c = \frac{3}{s} \tag{7}$$

The other model was

$$\left.\frac{\theta}{\theta_c}\right)_{model} = \frac{3(-s + 4.3)}{\left[s^2 + 2(0.18)(3.6)s + (3.6)^2\right]} \tag{8}$$

which resulted from closing the loop around

$$Y_p Y_c = \frac{3}{s}\frac{(-s + 4.3)}{(s + 4.3)} \doteq \frac{3}{s} e^{-0.47s} \tag{9}$$

Neither of these gave better results than the simple unity gain model.

Error Filter Characteristics

The original reason for providing an error filter was to have a direct method of emphasizing certain frequencies. However, changing the input characteristics has approximately the same effect. In fact, if the cost function consists only of model matching error, the input and filter dynamics are completely interchangeable. By properly choosing the input dynamics and cost function, good results were obtained without an error filter.

Proposed Problem Structure

The above paragraphs have discussed the various elements involved in the formulation of the SAS optimization problem. At this point we will summarize the proposed method of structuring the problem. This presentation will first consider situations in which the primary concern is manual control of attitude alone. Then extensions to manual path control will be discussed.

An example of manual attitude control is pitch control with the elevator. The proposed problem structure for this situation is shown in Fig. 2 for the case of only SAS feedbacks to the elevator. Later this form will be applied to the F-4 at three flight conditions. Note that because the model is assumed to be unity gain, the model matching error is the same as the manual tracking error, θ_e.

The form of the pilot describing function, Y_p, is not shown in Fig. 2. The recommended form is

$$Y_p = \frac{K_p}{0.1s + 1}\left(\frac{s - 16}{s + 16}\right)^2 \tag{10}$$

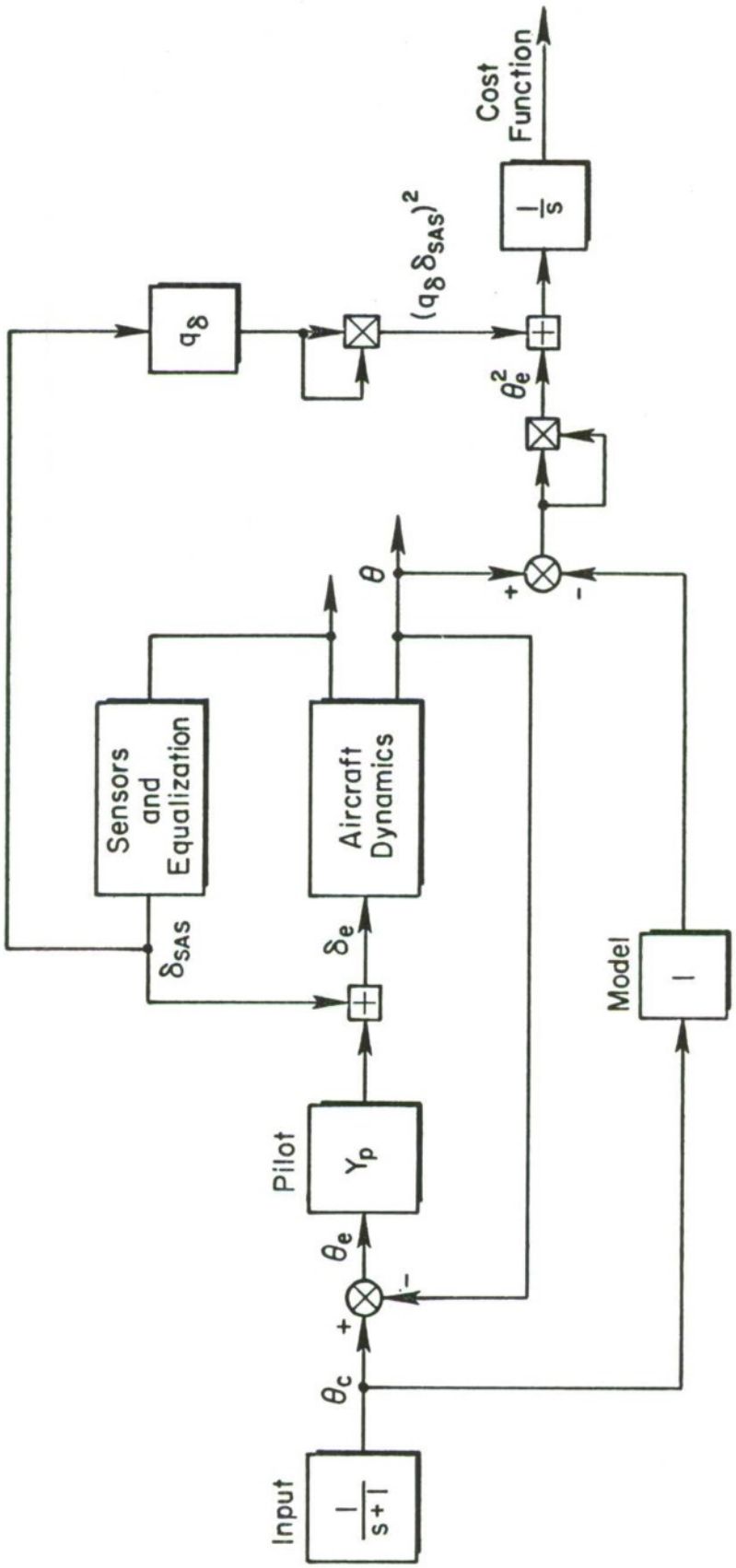


Figure 2. Proposed Problem Structure for Manual Pitch Control

The fixed elements represent typical pilot lags. The poles and zeros at 16 are an approximation to a 0.25 sec time delay. The gain, K_p, is a parameter to be optimized along with the SAS parameters.

It should be noted that the pilot model of Eq. 10 has no lead term. Lead was omitted because we found this improved the SAS characteristics in our example applications. In particular, including pilot lead led to a problem with the normal acceleration feedback to the elevator. In the low dynamic pressure examples, this feedback is needed to augment the short-period frequency. However, when lead was included in the pilot model the acceleration gain was too low and frequently had the wrong sign. Removing the lead in the pilot model allowed adequate augmentation of the short-period frequency.

The method of selecting the relative cost function weight on SAS control, q_δ, was described earlier. The weight is the magnitude of the pitch/elevator transfer function asymptote at 4 rad/sec.

In some situations, such as landing approach, it may be necessary to consider manual path control which is generally a multiloop task. Although we have not actually tried a problem of this type, we can estimate the modifications which would be required. First, the manual feedback structure would have to be expanded to include the outer path loop. The error term in the cost function should be changed to the outer loop or path error and the relative weight on SAS control changed accordingly. The relative weight on SAS control should be the path/elevator transfer function at roughly 1 rad/sec. Finally, it would probably be desirable to change the input characteristics. Path inputs and disturbances are generally lower bandwidth than attitude ones. Accordingly, an input time constant on the order of 2 sec, rather than 1 sec, is more appropriate.

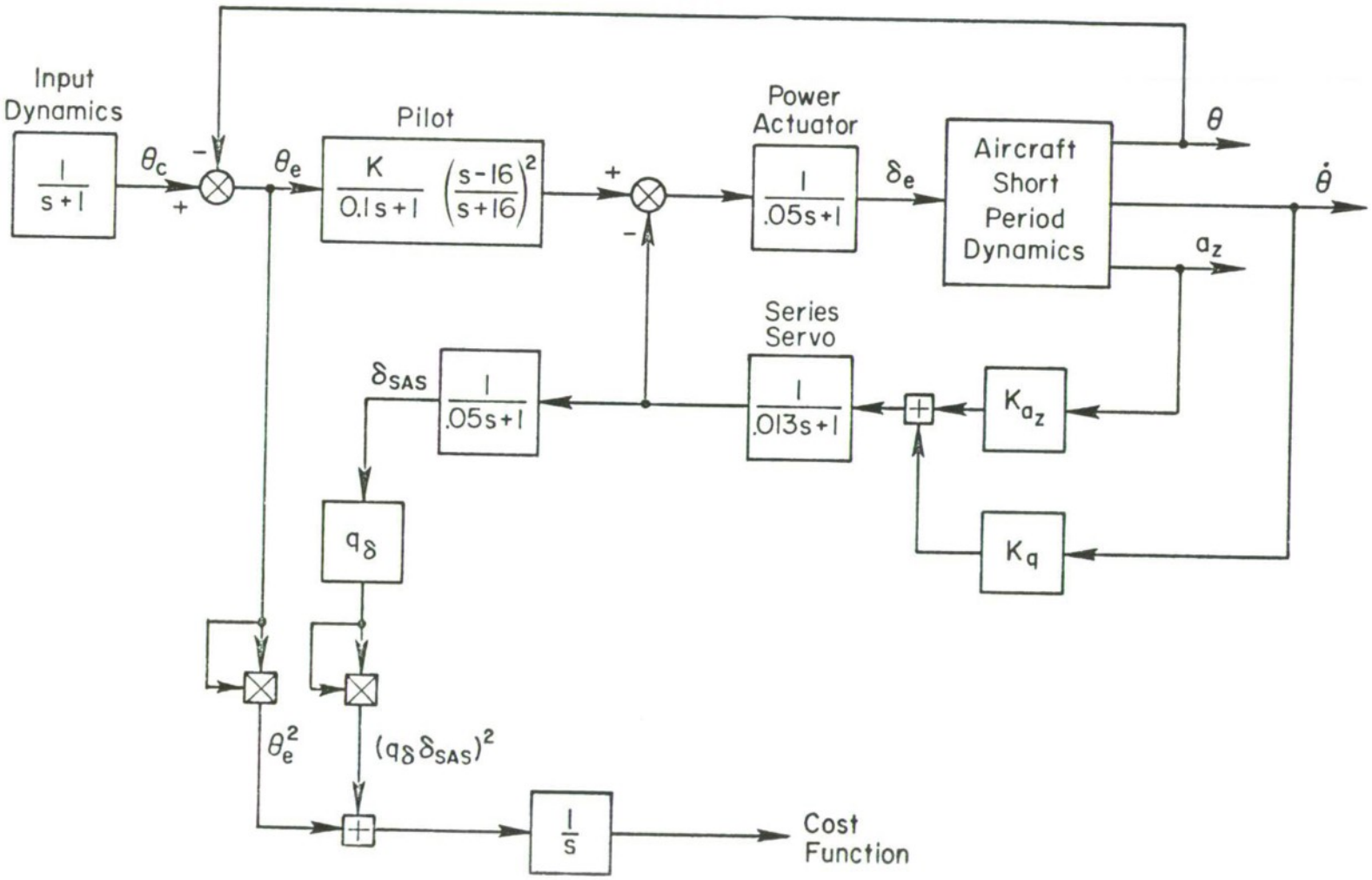
APPLICATION EXAMPLES

Example Problem

For an example application, the proposed design procedure was applied to the longitudinal axis of the F-4 airplane. Three different flight conditions were considered in the analysis. These correspond to a very low short-period frequency, a moderate value, and a relatively high frequency. The basic objectives of a SAS would be to augment the short-period frequency when it is low and to increase the short-period damping for the high-frequency case. To accomplish this the SAS was configured as normal acceleration and pitch rate feedbacks to the elevator. The accelerometer in all three cases was located at or very near the center of rotation for elevator inputs. The complete structure of the resulting optimization problem is shown in Fig. 3. The actuator dynamics are the same as those in the current F-4, Ref. 3. The vehicle stability derivatives were also taken from Ref. 3.

Results

The results of the three optimizations are summarized in Table II. Note that for the two subsonic flight conditions the feedbacks considerably augment



Free Parameters (to be optimized): K_{a_z} , K_q , K_p

Figure 3. Problem Structure for Examples

TABLE II

EXAMPLE OPTIMIZATION RESULTS

M	0.5	0.7	1.2
h (10^3 ft)	5	40	5
Accelerometer gain, K_{a_z} (rad/ft/sec^2)	−0.001168	−0.00703	-0.469×10^{-4}
Pitch rate gain, K_q (sec)	0.210	0.794	0.0636
Pilot gain, K_p	0.399	0.839	0.538
$[\zeta_{sp}, \omega_{sp}]$	[0.403 , 2.06]	[0.472 , 0.686]	[0.289 , 7.30]
Closed-loop poles,* accelerometer loop closed	[0.224 , 2.99](20.4)(76.9)	[0.060 , 2.42](20.3)(77.1)	[0.256 , 7.56](20.4)(77.0)
Closed-loop poles,* both SAS loops closed	[0.569 , 3.54](16.9)(77.7)	[0.752 , 3.00](14.9)(78.2)	[0.422 , 8.96](15.6)(78.1)
Closed-loop poles,* manual and SAS loops closed	(0.224)[0.290 , 3.28][0.827 , 11.9] [0.975 , 21.1](77.6)	(0.099)[0.376 , 2.93][0.824 , 11.7] [0.975 , 20.4](78.1)	(0.571)[0.256 , 5.76][0.549 , 13.9] [0.952 , 24.3](78.0)

*(p) denotes (s + p); $[\zeta , \omega]$ denotes $[s^2 + 2\zeta\omega s + \omega^2]$

the short-period frequency and damping. For the high dynamic pressure case the short-period frequency is already quite large and the optimal accelerometer gain is very low. In fact, this feedback has little effect on the short-period mode. The pitch rate feedback increases the damping and also significantly increases the frequency. The reason the pitch rate feedback increases the frequency is because of the significant actuator lags. A root locus plot of this feedback is shown in Fig. 4. Feedback gains higher than the optimal solution would increase the short-period damping but would also significantly increase the frequency, which is already quite high. The optimal solution is apparently a reasonable compromise between increasing the damping and not overly increasing the frequency.

A subjective judgment of the optimal solutions for these three cases is that the systems are at least reasonable and probably quite good. A comparison of the SAS/airplane characteristics with the military handling qualities specification is given below.

While the optimal solutions for the SAS feedbacks appear to be good, the results for the manual feedback are somewhat unusual. Asymptotic Bode plots of the manual loop closures are shown in Fig. 5. Admittedly these closures are not the type one would normally make in a pilot-vehicle analysis. The most drastic departures from the normal pilot model adjustment rules are the very low crossover frequencies. During our early investigations we did run across some cases in which the manual closures conform to the adjustment rules, that is, reasonably high crossover frequencies and K/s-like characteristics in the region of crossover. However, in all of these situations the SAS characteristics were very poor. In all cases the accelerometer feedback had the wrong sign, such that it reduced the short-period frequency rather than increased it, and, in addition, with both SAS loops closed the short-period mode was either overdamped or very nearly so. The objective of this project was to develop a SAS design procedure, not a technique for automated pilot-vehicle analysis. Therefore, given the choice between a problem formulation which gave good results on the SAS feedbacks or one which gave good results on the manual feedbacks, the choice was clear. The proposed procedure apparently has merit in terms of determining SAS characteristics even though the associated manual feedbacks are not realistic.

Comparison with Military Handling Qualities Specification

The optimal solutions were compared with the current military specification, Ref. 4. The comparison of the optimal results with the short-period requirements of the current military specification is shown in Table III. The results are easily within the specification limits.

SUMMARY

A systematic design procedure for stability augmentation systems has been presented. The two key features of this procedure are:

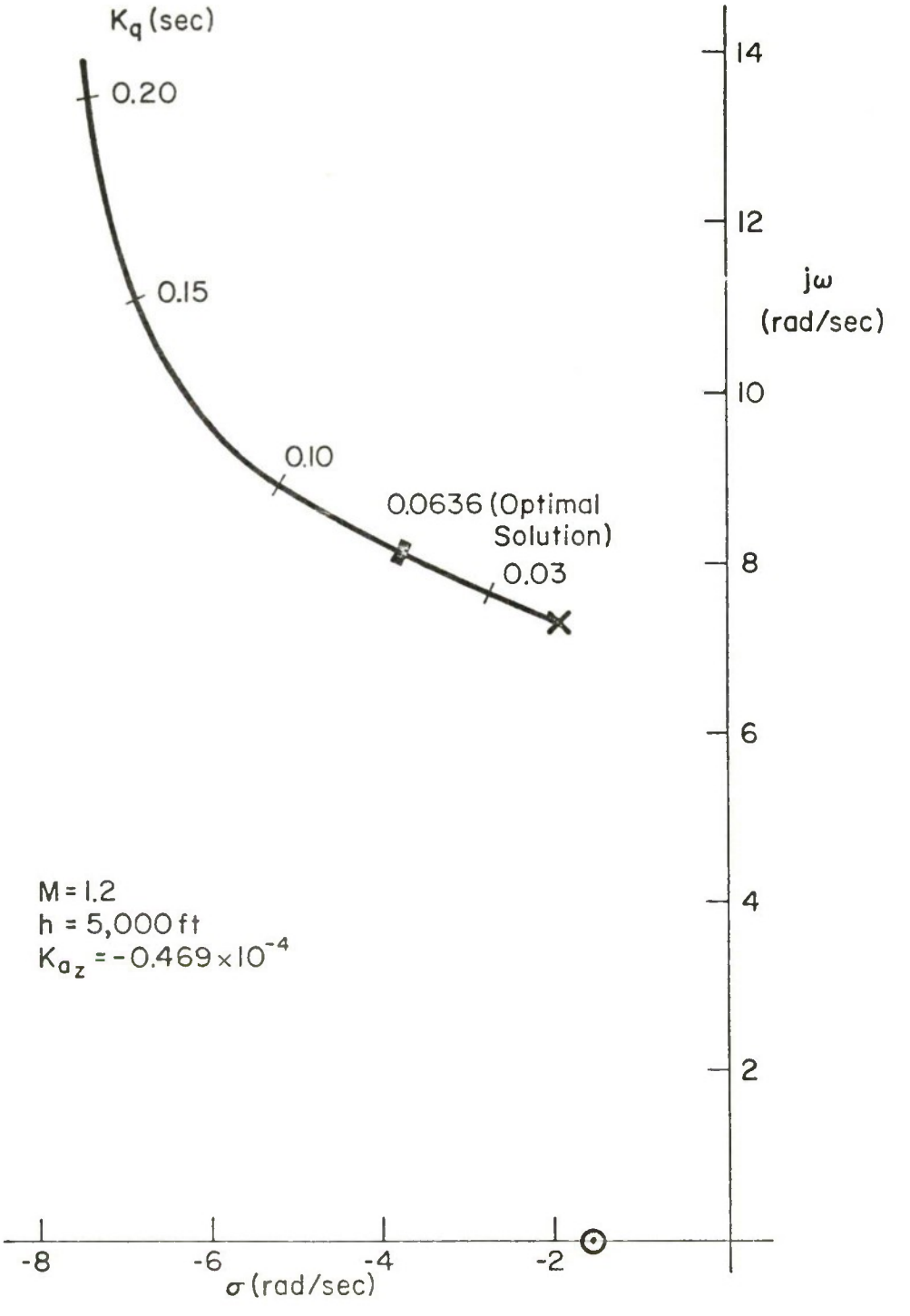


Figure 4. Pitch Damper Root Locus

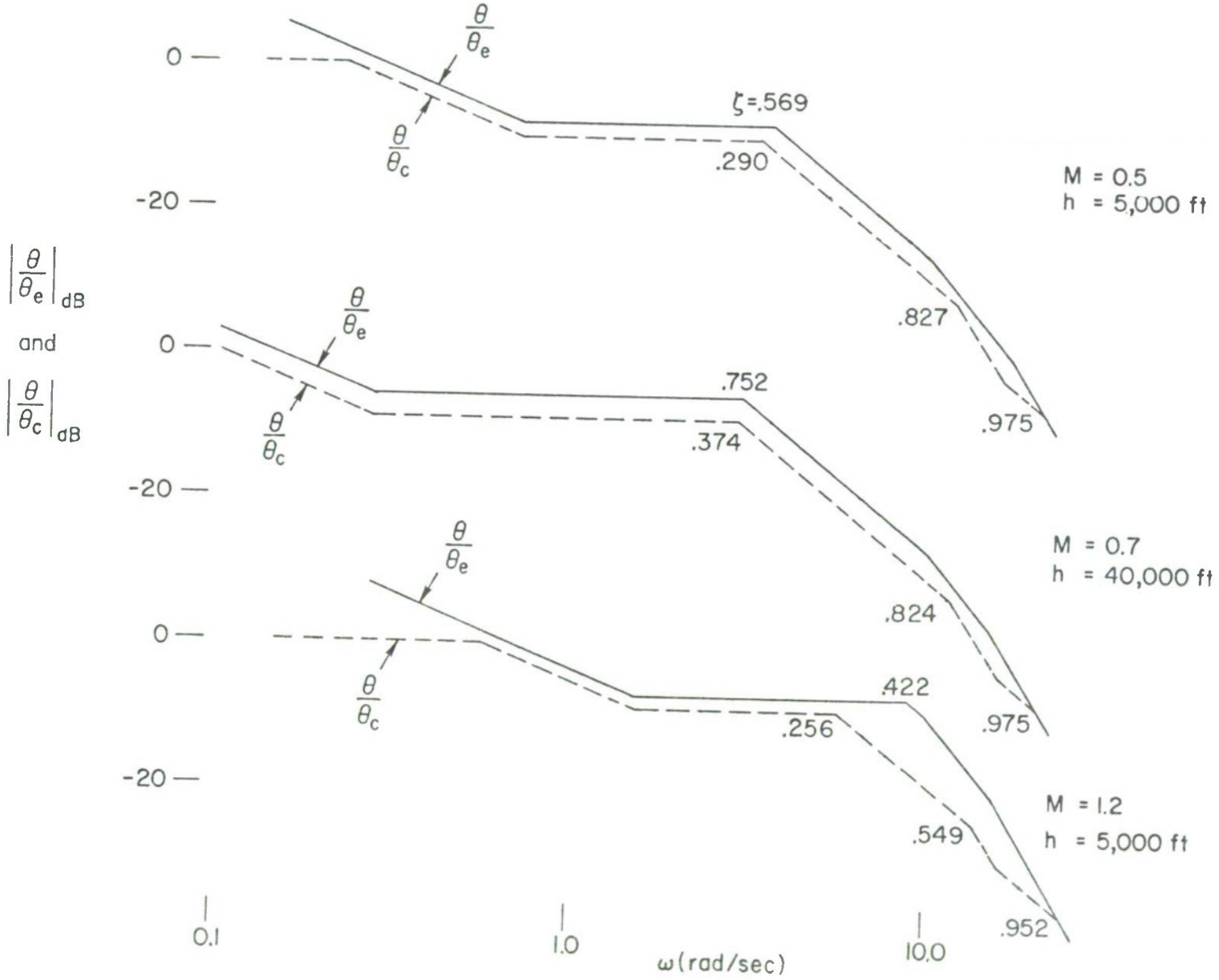


Figure 5. Manual Loop Closure Examples

TABLE III

COMPARISON OF OPTIMIZATION RESULTS WITH MILITARY SPECIFICATION

M	0.5	0.7	1.2
h (10^3 ft)	5	40	5
n/α (g's/rad)	11.75	5.79	62.5
ω_{sp} (rad/sec)			
Airplane alone	2.06	0.686	7.30
Airplane + SAS	3.54	3.00	8.96
Spec* limits	1.81 — 6.50	1.27 — 4.57	4.18 — 15.0
ζ_{sp}			
Airplane alone	0.403	0.472	0.289
Airplane + SAS	0.569	0.752	0.422
Spec* limits	0.35 — 1.30	0.35 — 1.30	0.35 — 1.30

*Reference 4 requirements, Category A, Level 1.

1. The selection of essential feedbacks from an examination of several key handling quality metrics

2. The use of parameter optimization techniques to determine the numerical values of the SAS parameters.

A two-stage parameter optimization is proposed. In the first stage, a variety of single case optimizations are performed, i.e., the parameters are optimized for each combination of flight condition and vehicle configuration. From these results, gain scheduling requirements are estimated. The second stage is the multicase optimizations where parameters (including gain scheduling parameters) are optimized over a number of combinations of flight condition and vehicle configuration.

The simple examples presented here have demonstrated the feasibility of this approach. Single case optimizations were performed for a longitudinal SAS for the F-4 aircraft at three very different flight conditions. All three used the same SAS form (pitch rate and normal acceleration feedback to the elevator), the identical problem formulation, and the same procedure for determining cost function weights. The resulting systems were judged quite satisfactory and, in fact, easily meet the short-period requirements of the military handling qualities specification, Ref. 4.

While the success achieved for these simple examples is highly encouraging, we recognize the likelihood of some unforeseen difficulties in other situations. A complete substantiation of the procedure will require applying it to:

1. Problems where manual outer loop, path control is important, such as landing approach

2. Situations requiring SAS feedbacks to more than just the elevator, such as to the throttle and direct lift devices

3. Lateral/directional SAS design

In addition, the multicase optimization step should be tested.

Successful implementation of this design procedure could greatly reduce the required time and cost for SAS design. It should also result in better SAS designs and thereby better handling qualities for future aircraft. With this great potential, additional research on this design procedure is strongly recommended, especially as regards the items mentioned in the previous paragraph.

REFERENCES

1. Stapleford, R. L., D. T. McRuer, L. G. Hofmann, and G. L. Teper, A Practical Optimization Design Procedure for Stability Augmentation Systems, Systems Technology, Inc., Tech. Rept. 187-1, Dec. 1969, (forthcoming AFFDL-TR-).

2. Hofmann, L. G., An Approach to Direct Design of Conventional Closed-Loop Controllers, Systems Technology, Inc., Paper 83, presented at SAE A-18 Committee Meeting, 12-14 Mar. 1969.

3. Bridges, B. C., Calculated Longitudinal Stability and Performance Characteristics of the F-4B/C/D/J and RF-4B/C Aircraft Plus the AN/ASA-32H Automatic Flight Control System, McDonnell-Douglas Rept. No. F934, 19 Apr. 1968.

4. Flying Qualities of Piloted Airplanes, MIL-F-008785A(USAF), 31 Oct. 1968.

A PILOT-VEHICLE SYSTEMS APPROACH TO LONGITUDINAL FLIGHT DIRECTOR DESIGN

R. H. Klein, D. T. McRuer, and D. H. Weir
Systems Technology, Inc.
Hawthorne, California

ABSTRACT

Recent developments in the theory of manual control displays now make feasible the statement of principles for analytical design of flight directors, given the dynamics of the (augmented) vehicle and its manual control system. The principal result from the theory is that there are effective director/vehicle controlled element dynamics which are preferred from the standpoint of pilot response and system performance. Other considerations include motion harmonization, display correspondence, and monitoring compatibility. This leads to rules and analytical procedures which allow the director computer feedbacks to be selected, weighted, and equalized to provide an effective director/vehicle system which satisfies both pilot-centered and guidance and control requirements. This paper summarizes the requirements and illustrates the analytical process for longitudinal control of transport-type aircraft during landing approach.

INTRODUCTION

This paper applies the existing "theory of manual control displays" (Refs. 1-5) to obtain design principles for advanced flight director systems and illustrates their application with a landing-approach example. The feedback control system aspects of the director/pilot/vehicle system are considered as a whole. In terms of hardware, this impacts primarily on the director computer because this is where the weighting and equalization of feedback signals occurs. Selecting the computer functions involves the following considerations:

- The system should possess adequate guidance and control properties.
- The dynamic characteristics of the effective vehicle/director controlled element should minimize the equalization and gain adjustment demands imposed on the pilot.
- The command signal should induce acceptable motion harmony when the pilot closes the loop.
- The displayed signals should be internally consistent and compatible with the real world.

The emphasis in this analytical development is on the longitudinal landing approach task — from beam acquisition to flare initiation. Nevertheless, the evolved design principles can be extended to other tasks such as attitude hold, altitude hold, flare, lateral path following, etc.

---

*Accomplished in part under NASA Contract NAS2-3746.

The paper begins with a brief background note. Then, two kinds of requirements for director design are presented; those relating to basic guidance and control considerations, and additional ones which result from the presence of a human pilot in the control loop. Following the requirements, two analytical examples illustrate their implementation in the design of the director computer. The first example is a basic director (containing only pitch attitude and beam deviation) which satisfies some of the requirements. Its deficiencies are used to motivate the synthesis of an advanced director (containing properly equalized additional feedbacks) which best satisfies both types of requirements.

BACKGROUND

A flight director system combines display, computation, pilot, and effective (augmented) vehicle in a feedback control system, as shown in Fig. 1 for approach.

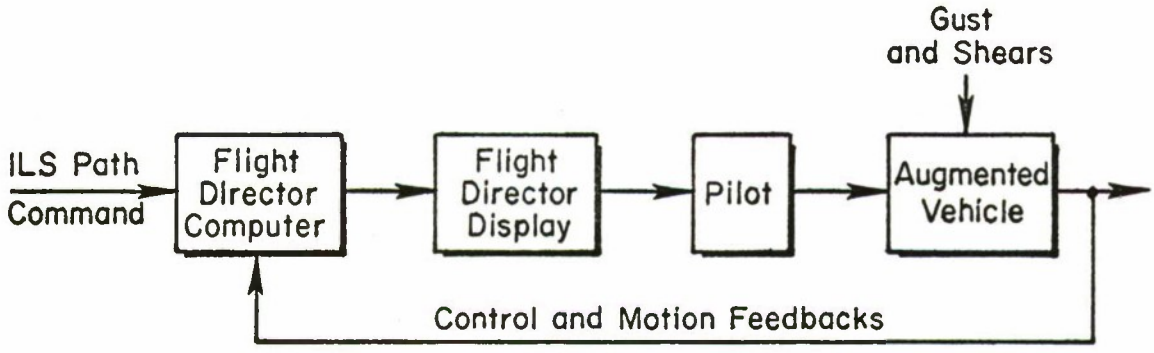


Figure 1. Basic Flight Director System Elements in Approach

The command elements in the director display provide lateral and vertical steering signals made up of a combination of desired path and aircraft motion quantities. These are shaped, filtered and mixed appropriately to permit the pilot to close the flight director system loop with ease and efficiency. The status information includes an artificial horizon for all purpose use, and other pictorial information pertinent to a particular phase of flight. In landing approach, localizer and glide path signals are invariably presented, and indications of altitude and airspeed error are provided on the more modern instruments.

The nub of the dynamic design problem is the selection of the appropriate mix of signals to make up the steering commands. Historically, this has been done in two general ways:

- By adapting and displaying the output of an automatic control system
- By mechanizing the director computer based on guidance and control requirements, and adjusting the various feedback gains during simulation and flight test for acceptable pilot opinion and overall system performance.

Both satisfy the guidance and control requirements and the first is useful for monitoring, but neither pays explicit attention to the human pilot's needs until the system is tested or simulated. Time and money can be saved by considering the pilot properties at the outset, and designing the director system to:

- Reduce pilot remnant (unwanted control action) by reducing scanning and the need for pilot equalization
- Reduce pilot equalization and required gain variation
- Provide for a wide range of pilot gain to permit either very loose or very tight control with good characteristics for both

All lead to superior control. Thus, advanced director design must be based on two sets of requirements which are

- Fundamental; independent of whether the controller is an automatic or human pilot; and
- Human centered

These requirements are developed in the next section, followed by an example of the analytical synthesis procedure.

GUIDANCE AND CONTROL REQUIREMENTS

The guidance and control requirements are to establish the aircraft on glide path, and to reduce any path errors to zero in a stable, well-damped and rapid manner. They lead to outer loop feedbacks which are required to accomplish the mission, plus inner loop feedbacks which permit the first to function. The basic system for longitudinal control is shown in Fig. 2.

The fundamental path quantity in the block diagram is the beam deviation, d_e; this is corrupted from the ideal by beam bends and by noise in the airborne equipment. The actual physical signal is a glide path error angle, γ_e, converted from the deviation by the decreasing range, R. The path error and the inner loop feedbacks are combined in the director computer and displayed to the pilot on the director indicator. The pilot can conceivably close other loops using raw data from the instrument panel, but these are unnecessary if the director is properly designed. To eliminate the need for pilot gain reduction with decreasing range, the time-varying gain G_{γ_e} is multiplied by the range to obtain a constant coefficient operator $G_{d_e} = G_{\gamma_e} R$.

Inputs to the system of Fig. 2 may include the following:

- Step (offset) glide slope command, including initial beam acquisition
- Dual angle beam, representing a ramp change from one beam angle to another

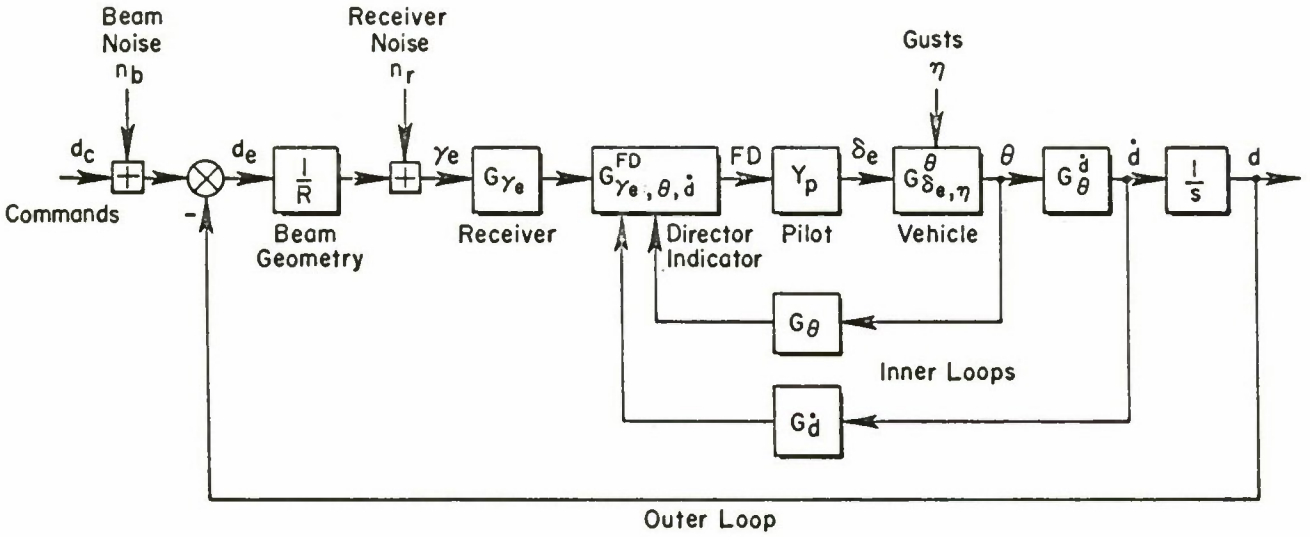


Figure 2. Block Diagram for Approach Control with Flight Director

- Random signals, such as beam bends and receiver noise
- Trim changes in the vehicle
- Discrete changes in the horizontal and vertical winds, including increasing or decreasing headwinds and tailwinds (shears)
- Horizontal and vertical random gusts

The adequacy of the system is determined by the path errors in the presence of these inputs.

Table I provides a summary at the outset of feedbacks in the computed director signal which can satisfy the fundamental requirements for the system functions which must be performed. These are developed in Refs. 6 and 7, for example, and they are elaborated in the example analysis, below. Beam deviation provides the basic outer loop for command following, and its gain determines the bandwidth or stiffness of the system. Damping is achieved by feeding back functions of the attitude, rate of climb, and/or beam rate. Pitch attitude also has a primary function in maintaining attitude-stability and avoiding overrotations. Windproofing is accomplished by adding various functions of beam deviation. Integral of beam deviation is included to reduce path errors in the presence of very low frequency beam commands or wind shears.

PILOT RELATED REQUIREMENTS

The presence of a human pilot in the control loop requires a division of functions between the pilot and the director computer, and adds another dimension to system performance considerations. The feedbacks must be

TABLE I

SUMMARY OF FEEDBACKS TO SATISFY FUNDAMENTAL REQUIREMENTS

SYSTEM FUNCTIONS — FUNDAMENTAL REQUIREMENTS	FLIGHT DIRECTOR FEEDBACKS
Path command and stiffening	Beam deviation, d
Short-period attitude regulation	Attitude, θ, at short-period frequencies
Path damping	Attitude, θ, at path frequencies or beam rate, $\dot{d}$
Short-period damping	Attitude rate, $\dot{\theta}$
Mid-frequency windproofing	Beam rate, $\dot{d}$
High frequency windproofing	Vertical acceleration, a_z
Curved path following	Beam double integration, $\iint d\,dt\,dt$
Path angle trimming	Beam integration, $\int d\,dt$
Low frequency windproofing	Beam integration, $\int d\,dt$

selected, equalized and weighted, not only to obtain good overall performance, but also to be compatible with good subjective pilot ratings. The pilot-centered requirements are elaborated below.

Equalization for Minimum Pilot Effort

It is well known that the human pilot adapts his characteristics to compensate for dynamic deficiencies of the effective controlled element. As part of this adaptation, he may be forced to develop low-frequency lead or to adjust his gain precisely. When low-frequency lead is required, a cost in pilot dynamic capacity is incurred (Refs. 8-10) which is reflected in increased pilot time delay and remnant. These cause a deterioration in system performance, reduce the additional cockpit workload which can be handled, and lead to a degradation in pilot ratings.

As a result of these human pilot properties, the effective control element should be constructed to:

- Require no low-frequency lead equalization
- Permit pilot loop closure using a wide range of pilot gains

This can be achieved when the effective controlled element approximates either a gain, K, or an integration, K/s, in the frequency range of pilot/director/vehicle system crossover. Although the pilot's effective time delay is somewhat less with a simple gain than an integrator the aircraft attitude in the frequency region about the short period is not proportional to elevator position, and it is not feasible to attain this without additional automatic feedbacks. An alternative is to shape the ILS signal with a large lead (differentiator), but this would result in an unacceptable amplification of glide slope noise. Another possibility is to include a very high gain elevator feedback to the director, but this is unsuitable because the large feedback gain needed requires a reduction in display scale, thereby making the desired command inputs barely perceptible to the pilot.

A K/s effective controlled element is nearly as good as a pure gain from the standpoint of pilot response and performance in director-like tasks. With K/s the pilot transfer characteristics are approximately a gain plus time delay in the frequency region of control, his time delay will be close to minimum, and remnant can be minimized with the proper choice of controlled element gain. Lead generation requirements are small, although the pilot can use a small amount at high-frequency to reduce his effective time delay in the loop. This lead can be minimized by making the controlled element more like a gain at high frequency.

Other requirements based on minimizing pilot effort are to:

- Range compensate the beam error so that the display/controlled element dynamics are approximately time invariant. The pilot can adjust to nonstationary situations, but it involves adaptation and learning which increases task difficulty and degrades performance.

- Account for other pilot workload and for unattended operation by providing dynamics which permit wide variations in pilot gain while retaining adequate gain and phase margins throughout the mid-frequency region. This implies that conditionally stable systems, and feedback of beam integral are undesirable.

These minimum pilot effort requirements must be further tempered with the considerations developed below.

Motion Harmony

Motion harmony considers the ways in which the various motions of the aircraft interrelate and how this affects the pilot. In essence, the airframe motions generated by the pilot in attempting to reduce the director error signal to zero should be similar to those which he experiences under other manual control conditions. This is desirable both for the pilot's internal self-monitoring functions and for the monitoring of pilot activity by the copilot using the full instrument panel. Quantitatively, this requirement is assessed using modal response ratios as shown subsequently.

Another requirement related to motion harmony is the compatibility between director and autopilot during an automatic approach. This might be taken to mean that the autopilot and director feedbacks should be identical, but certain signals such as the beam integral are more appropriate to the automatic system than to the director. These minor differences need not be significant since the fundamental goal is to make the autopilot and flight director correspond only at the dominant modes.

Face Validity and Display Consistency

Some elements of the director display are intended as instrument reproductions of those portions of the external world which are sources of visual flight cues. If the resulting abstraction evokes responses while on IFR that are similar to those under VFR conditions the display is adequate from a behavioral standpoint, and it may even be superior to VFR by providing cues which are unavailable from the visual scene. In a director these cues are command signals which the pilot is to follow, and the rest of the display presents status information. If the status information corresponds to the actual situation, it has a high degree of "face validity." For example, the artificial horizon corresponds directly with the actual horizon. Other status elements are the glide slope and localizer signals.

The command signals must also have face validity in some sense. In this case the cue is different from status information, because the command signal is a mixture of control and vehicle motions rather than an abstraction of a real world cue. However, some correspondence does exist between the command signal and the dominant vehicle motion or control quantity in each of several frequency bands. Thus the command signal may be said to have some degree of consistency with the status elements on the display. This is necessary if the pilot is to believe his instruments. For example, the flight director should be dominated by attitude at short period frequencies but retain only beam deviation at low frequencies to avoid a glide path standoff error.

Director computer requirements derived from consideration of equalization to minimize pilot effort, motion harmony, and face validity and display consistency are summarized in Table II. Taken together, the fundamental and

TABLE II

FLIGHT DIRECTOR COMPUTER FUNCTIONS
FOR HUMAN-CENTERED REQUIREMENTS

REQUIREMENT	FUNCTION
K/s effective controlled element	$\dot{h}$ or $\dot{d}$ feedback at mid-frequencies
Display consistency	Only d feedback at very low frequency. θ feedback at short period frequencies.
Motion harmony	Response with flight director similar to that for raw data (or VFR); and similar to that for autpilot, i.e., θ and $\dot{h}$ inner loop feedbacks.

pilot-centered requirements prescribe the flight director computer feedbacks, as well as the general nature of their weighting and equalization, needed to accomplish a landing approach in the presence of disturbances. The implementation and analytical interpretation of these requirements for a conventional transport-type aircraft will now be presented.

SYNTHESIS OF AN ADVANCED LONGITUDINAL DIRECTOR

The derivation of the director/vehicle controlled element is accomplished by setting up a system which has feedbacks that satisfy the functional requirements in a minimal way, and then subjecting it to constructive criticism. The minimum systems are derived in Ref. 7 on the basis of the steady-state path following ability in the presence of the following inputs:

- Step offset from the beam
- Change in beam angle
- Curved beam
- Step changes in wind velocity
- Ramp changes in wind velocity (shears)

The result is a minimum system for each type of input, and the one that provides some gust regulation is beam deviation plus washed-out pitch attitude. These feedbacks are also the minimum acceptable from the other requirements of Tables I and II.

The deficiencies of this basic deviation/attitude system provide the start for evolving an advanced director that has superior regulation and command following properties, and which better satisfies the pilot's subjective feeling for responsiveness, validity, and motion harmony. This is done by an analytical development using the DC-8 transport as a numerical example.

Axis System and Vehicle Equations

The analysis uses a conventional body-fixed stability axis system. For landing approach the unperturbed x-axis is aligned with the glide slope, making $\theta_o = \gamma_o$. When the aircraft's attitude is perturbed, the orientation of the body-fixed x- and z-axes becomes that shown in Fig. 3.

The longitudinal dynamics of the vehicle are assumed to be described by the linearized three degree of freedom perturbation equations given in Eq. 1. Values for the airframe parameters and stability derivatives for the DC-8 jet transport example are contained in Ref. 7. Numerical values for the transfer function numerators and denominators to be used in the example are given in Table III.

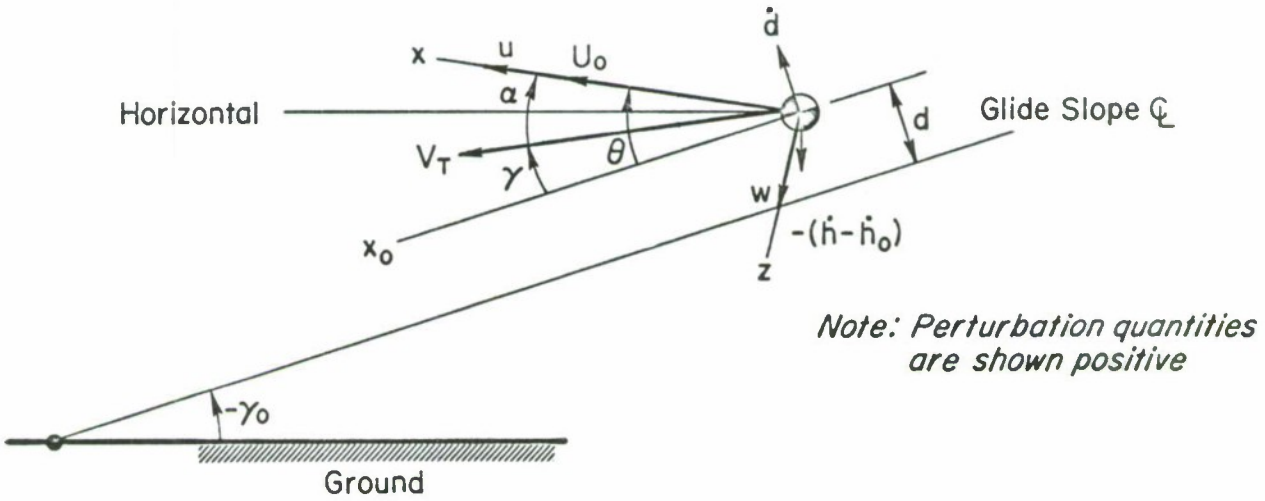


Figure 3. Body Axis System

$$\begin{bmatrix} s - X_u & -X_w & g\cos\Theta_o \\ -Z_u & s - Z_w & -U_o s + g\sin\Theta_o \\ -M_u & -(M_{\dot{w}}s + M_w) & s(s - M_q) \end{bmatrix} \begin{bmatrix} u \\ w \\ \theta \end{bmatrix} = \begin{bmatrix} X_{\delta_e} \\ Z_{\delta_e} \\ M_{\delta_e} \end{bmatrix} \left[\delta_e\right] \qquad (1)$$

$$\dot{d} = -w + U_o\theta + \ell_x\dot{\theta} \text{ , at station } \ell_x$$

$$\dot{h} = -w\cos\Theta_o + u\sin\Theta_o + U_o\cos\Theta_o\theta$$

$$a_z = \dot{w} - U_o\dot{\theta} + g\sin\Theta_o\theta - \ell_x\ddot{\theta}, \text{ at station } \ell_x$$

TABLE III

LONGITUDINAL STABILITY AXIS TRANSFER FUNCTIONS
FOR THE DC-8 IN THE LANDING APPROACH CONFIGURATION

Denominator, Δ	[0.0865 ; 0.166][0.627 ; 1.23]*
Pitch Numerator, $N^{\theta}_{\delta_e}$	−0.915(0.101)(0.646)
Altitude Rate Numerator, $N^{\dot{h}}_{\delta_e}$	9.25(−3.63)(0.0352)(4.42)

*To simplify the notation, $A[s^2 + 2\zeta\omega s + \omega^2]$ is written $A[\zeta\ ;\ \omega)$ and $A(s + a)$ is written $A(a)$.

The Basic Beam-Deviation-Plus-Pitch-Attitude System

The beam deviation/attitude system is shown in the block diagram of Fig. 4, where G_h^{FD} and G_θ^{FD} are the respective feedback functions.

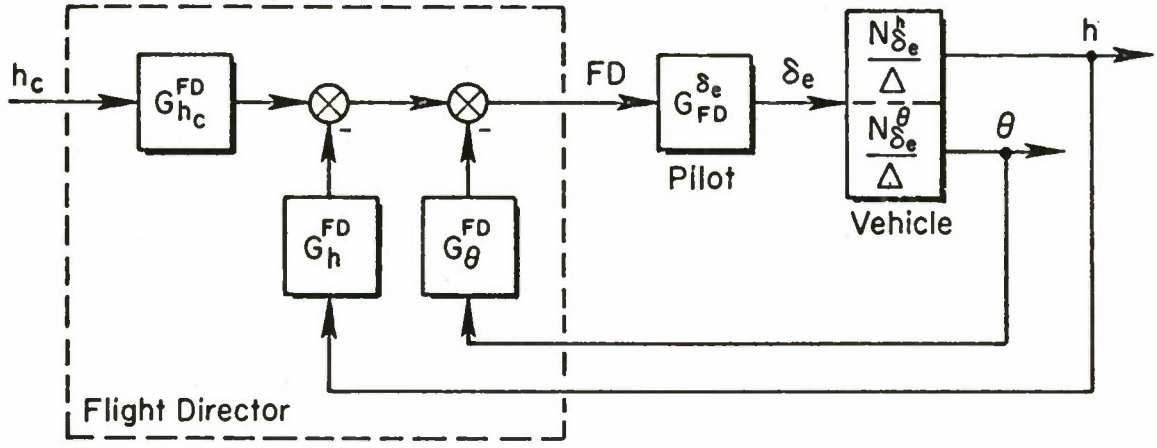


Figure 4. Simplified Director/Vehicle Block Diagram

The system has been made constant-coefficient by removing the range variation. Beam deviation, d, has been simplified to perturbation altitude, h, by assuming no steady-state wind effects to change the ground speed. Without wind, the small difference between d and h due to sin Θ_o is negligible for glide path angles of current ILS systems.

A central implication of the pilot-centered requirements is that the gains and equalizations be selected so that the net transfer characteristics from pilot elevator output to director instrument displacement look approximately like an integration, K/s. The analytical procedure for assessing and establishing this result is to add the component vehicle transfer functions with their associated equalization. The effective flight director transfer function is then,

$$\frac{FD}{\delta_e} = \frac{G_\theta^{FD} N_{\delta_e}^\theta + G_h^{FD} N_{\delta_e}^h}{\Delta} \qquad (2)$$

For the basic system the feedback functions are:

$$G_\theta^{FD} = \frac{K_\theta s}{s + 1/T_{WO}} \qquad (3)$$

$$G_h^{FD} = K_h$$

The controlled element transfer function is given by:

$$\frac{FD}{\delta_e} = \frac{\frac{K_\theta A_\theta s}{s + 1/T_{WO}}\left(s + \frac{1}{T_{\theta_1}}\right)\left(s + \frac{1}{T_{\theta_2}}\right) + K_h \frac{A_h^{\cdot}}{s}\left(s + \frac{1}{T_{h_1}}\right)\left(s + \frac{1}{T_{h_2}}\right)\left(s + \frac{1}{T_{h_3}}\right)}{\Delta} \tag{4}$$

The washout time constant is chosen to retain a good attitude signal at and below mid-frequency to ensure good path damping, but yet not too slow for very low frequency windproofing. In order to obtain a closed-loop system that will have the required path damping as well as a closed-loop path mode frequency greater than the phugoid, the washout inverse time constant must be less than the phugoid, ω_p. With approach speeds on the order of 200 to 300 fps, the washout time constant will generally be around 10 sec.

With a very slow washout the numerator of Eq. 4 combines into a low frequency dipole pair $(s + 1/T_{WO})'/(s + 1/T_{WO})$ which cancel to a first approximation, a first-order root at nearly $1/T_{h_1}$, and a second-order pair at an undamped natural frequency, ω_θ, determined by the $\sqrt{K_h/K_\theta}$ gain ratio. The approximate transfer function is:

$$\frac{FD}{\delta_e} \doteq \frac{K_\theta A_\theta \left(s + \frac{1}{T_{h_1'}}\right)\left[s^2 + \frac{1}{T_{\theta_2}} s + \frac{K_h A_h^{\cdot}}{K_\theta A_\theta T_{h_2} T_{h_3}}\right]}{s[s^2 + 2\zeta_p \omega_p s + \omega_p^2][s^2 + 2\zeta_{sp}\omega_{sp} s + \omega_{sp}^2]} \tag{5}$$

Figure 5 contains frequency response ($j\omega$-Bode) and root locus plots of this basic system transfer function for two values of the gain ratio, K_h/K_θ. The smaller value is given by the dashed line. The location of the ω_θ zeros is determined by the ratio $\sqrt{K_h/K_\theta}$. Note that at larger K_h/K_θ values (solid line) the system is conditionally stable and has no region of K/s-like amplitude ratio. This will make the system more sensitive to variations in pilot gain, and will restrict the pilot-vehicle system crossover to frequencies outside the crosshatched unstable region. The system becomes stable over a broad region as ω_θ is decreased. Also, as ω_θ is decreased the spread between ω_θ and ω_{sp} increases and a small region of K/s-like amplitude ratio is produced in between. As such, the form of the response is insensitive to changing gain and the bandwidth is proportional to the gain selected, while with K/s^2-like systems the closed-loop dynamics change sharply as the gain varies. The predominant K/s^2 nature of the effective controlled element in the anticipated crossover frequency region between the phugoid, ω_p, and short period, ω_{sp}, is the most apparent drawback of the basic flight director system.

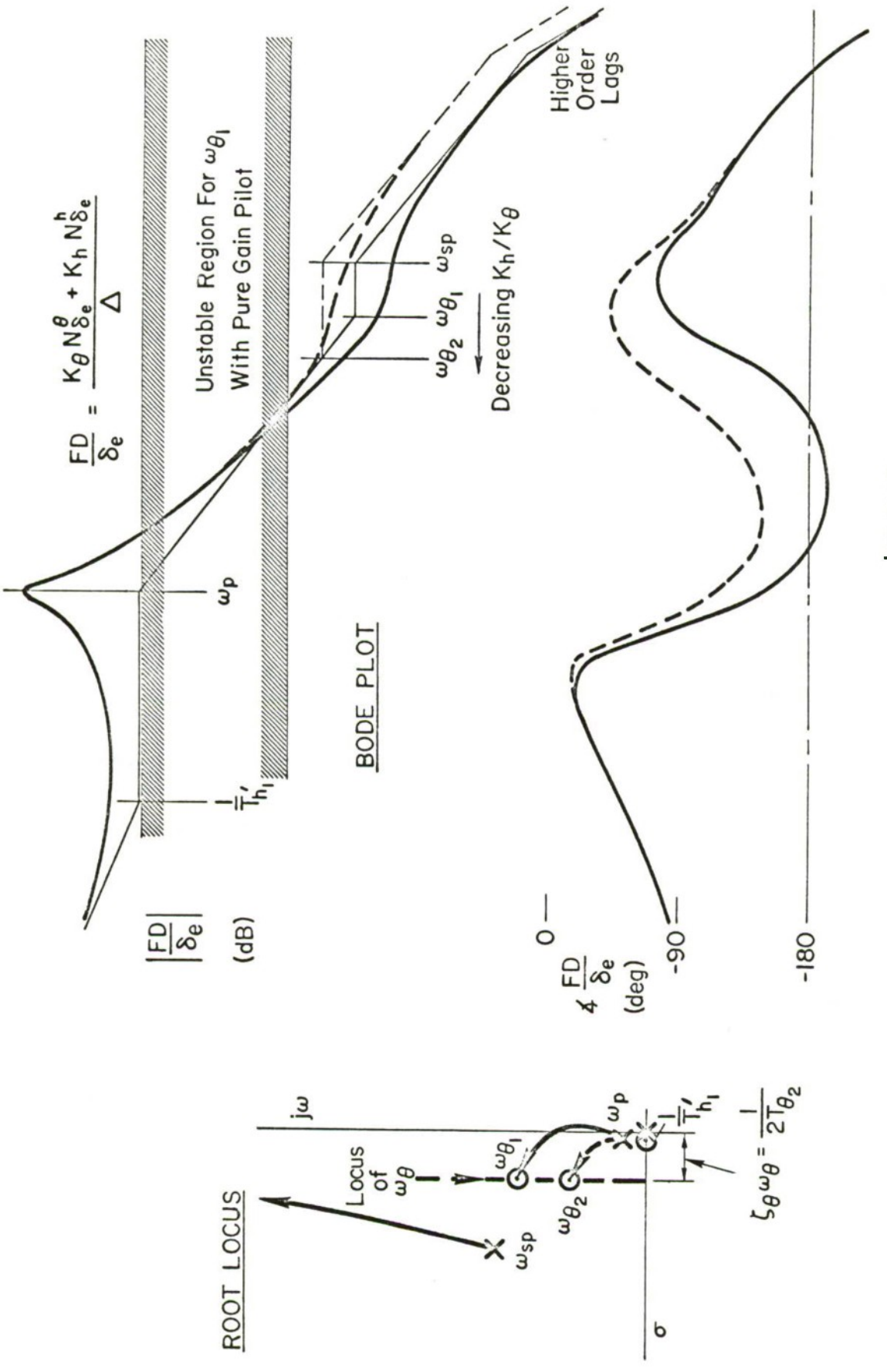


Figure 5. Variation in Director/Vehicle Properties with K_h/K_θ Weighting

Referring to Fig. 5 and Eq. 5, it can be seen that with large enough pitch gain, K_θ, the second-order ω_θ zeros can be overdamped to produce two first order zeros. This appears to improve the mid-frequency gain and produce the desirable K/s region, but it has limitations. The total numerator damping is constant so the two first orders can not be placed separately in the most desirable locations. For large θ gain the flight director will look very much like an amplified pitch attitude display, and be too busy in turbulence, violating display consistency. Attempts by the pilot to follow the bar may result in unacceptable normal accelerations and pitch attitude excursions — incompatible with good motion harmony.

Other overall deficiencies of the basic $\theta + h$ flight director include poor gust regulation due to the slow attitude washout, high sensitivity of the numerator zeros to slight changes in the K_h/K_θ ratio, and the K/s^2 amplitude ratio above the short period. This implies a need for pilot (or other) lead equalization in this frequency range, in order to extend the K/s region. Usually, this will not be a strong requirement, unless ω_{sp} is smaller than about 1 rad/sec. For lower short period frequencies, additional equalization in the director should be considered.

The advantages and deficiencies of the attitude/beam deviation system are summarized in Table IV.

TABLE IV

RELATIVE PROPERTIES OF THE BASIC DEVIATION/ATTITUDE DIRECTOR SYSTEM

ADVANTAGES	DEFICIENCIES
Simple to mechanize Provides display consistency with washed out attitude feedback	K/s^2-like amplitude ratio at mid-frequency when K_h/K_θ weighting is acceptable Poor w-gust windproofing due to slow washout K/s^2-like amplitude ratio at high frequency Maximum crossover frequency restricted by non-pilot lags in forward loop

Addition of Beam Rate Feedback to the Basic Director

Combining altitude rate or beam rate with attitude and beam deviation feedbacks provides more precise beam following, and better fulfillment of pilot-centered requirements. The basic effect is to change the second-order zeros in the flight director transfer function of Eq. 5 to

$$\left[s^2 + \left(\frac{1}{T_{\theta_2}} + \frac{K_{\dot{h}}}{K_\theta}\frac{A_{\dot{h}}^*}{A_\theta}\right)s + \frac{K_h}{K_\theta}\frac{A_{\dot{h}}^*}{A_\theta}\right] \tag{6}$$

Again, $1/T_{h_2'}$ and $1/T_{h_3}$ are assumed large relative to ω_{sp}. This quadratic may be separated into two independent first-order zeros. Each can be located independently to maximize the K/s region. This means placing them near ω_p and ω_{sp}, respectively. When the two roots are greatly different, Eq. 6 separates into

$$\frac{1}{T_1} = \frac{\frac{K_h Z_\alpha}{K_\theta}}{\left(\frac{1}{T_{\theta_2}} + \frac{K_{\dot{h}} Z_\alpha}{K_\theta}\right)} \doteq \frac{\frac{K_h Z_\alpha}{K_\theta}}{\frac{K_{\dot{h}} Z_\alpha}{K_\theta}\left[1 - \frac{K_\theta}{U_o K_{\dot{h}}}\right]} \doteq \frac{K_h}{K_{\dot{h}}} \tag{7}$$

and one large root,

$$\frac{1}{T_2} = -\frac{K_{\dot{h}} Z_\alpha}{K_\theta}\left[1 - \frac{K_\theta}{U_o K_{\dot{h}}}\right] \doteq -\frac{K_{\dot{h}} Z_\alpha}{K_\theta} \tag{8}$$

Thus, a combination of feedback gains can be selected for optimum zero location.

Figure 6 presents frequency response and root locus plots for the modified director/vehicle controlled element. The zeros have been selected as follows:

$$\frac{1}{T_1} \doteq \frac{K_h}{K_{\dot{h}}} \doteq \omega_p \tag{9}$$

$$\frac{1}{T_2} \doteq -\frac{K_{\dot{h}} Z_\alpha}{K_\theta} \doteq \omega_{sp} \tag{10}$$

The frequency response shows a broad K/s region between ω_p and ω_{sp}, with very little phase dip near ω_p. The path damping is now coming from the low frequency zero, $1/T_1$, due to the $\dot{h}$ feedback. Good high frequency properties are provided by the other zero, $1/T_2$.

The gust regulation and low frequency windproofing are also improved with the addition of $\dot{h}$ because the attitude feedback can be washed out much faster than in the basic flight director case without compromising the mid-frequency

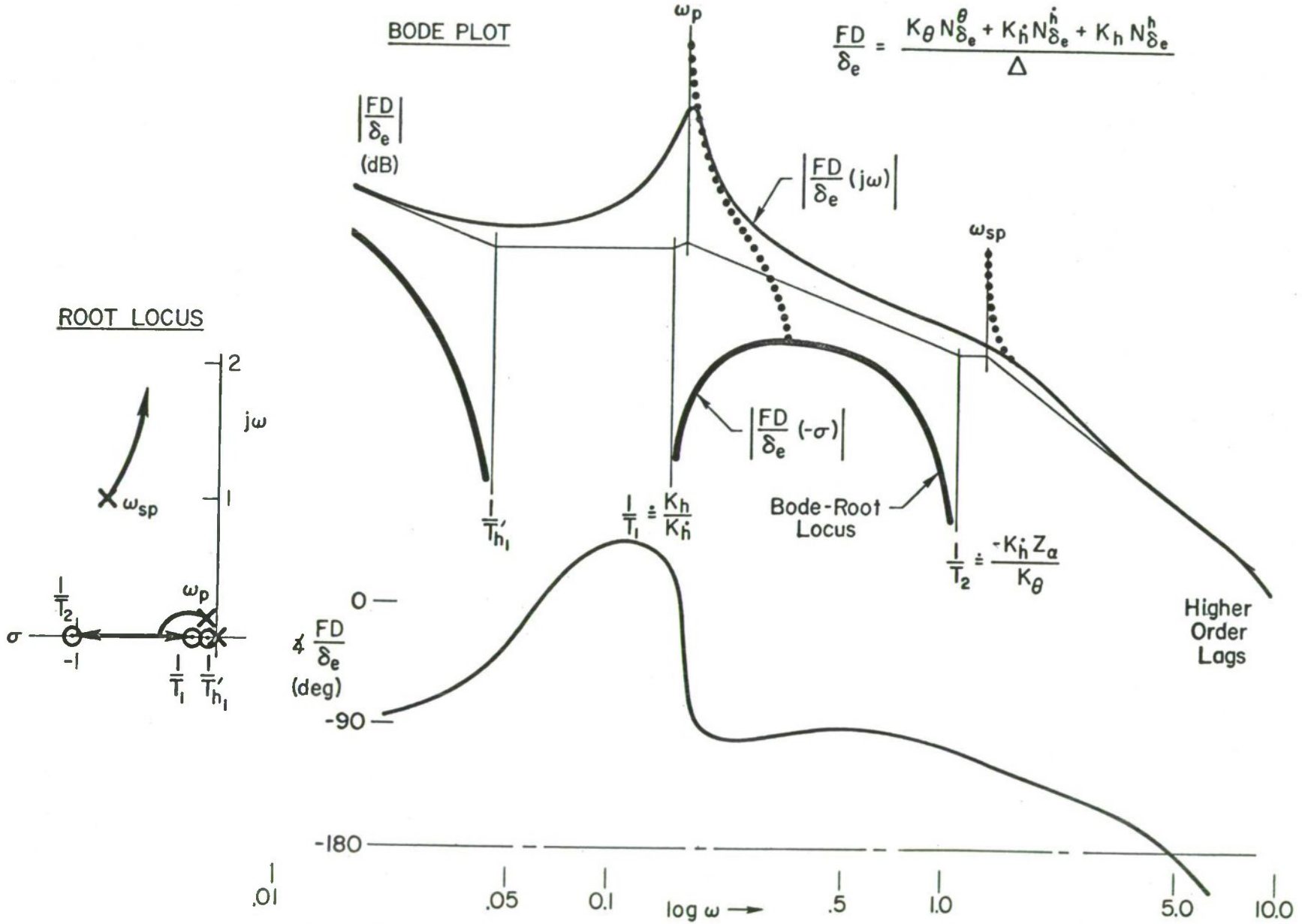


Figure 6. Director/Vehicle Properties with Beam Rate Added

path damping. The relationship between θ and $\dot{h}$ is helpful in determining the slowest reasonable pitch attitude washout time constant. In the low to mid-frequency region, a good approximation relating $\dot{h}$ and θ for elevator inputs is given by

$$\frac{\dot{h}}{\theta} = \frac{U_o}{T_{\theta_2}s + 1} \tag{11}$$

Thus, for frequencies below $1/T_{\theta_2}$, $\dot{h}$ and θ feedbacks are redundant (in the absence of winds), so θ can be washed out with a time constant of at least T_{θ_2}. Because the pitch attitude feedback provides the required attitude stability, the ultimate lower limit on the washout time constant is near the short-period frequency.

The only deficiency with the combined $\theta + h + \dot{h}$ flight director system is that the K/s region does not extend beyond the short-period frequency. This means that potential high gain pilot closures will require pilot lead equalization in the vicinity of the short period.

Addition of Pitch Rate to the $\theta + h + \dot{h}$ Director

Pitch rate can be included in the G_θ^{FD} control path to independently control the short-period damping ratio. The addition of pitch rate creates an additional zero in the flight director transfer function, FD/δ_e. Placing the zero near the short period makes the flight director transfer function K/s-like at and above ω_{sp}. The closed-loop short-period damping ratio will also increase as the pilot increases his gain.

The resulting feedback functions are then:

$$G_\theta^{FD} = \frac{K_\theta s}{\left(s + \frac{1}{T_{WO}}\right)} + K_{\dot{\theta}} s = \frac{K_{\dot{\theta}} s \left[s + \left(\frac{1}{T_{WO}} + \frac{K_\theta}{K_{\dot{\theta}}}\right)\right]}{\left(s + \frac{1}{T_{WO}}\right)} \tag{12}$$

$$G_h^{FD} = K_h + K_{\dot{h}} s = K_{\dot{h}}\left(s + \frac{K_h}{K_{\dot{h}}}\right) \tag{13}$$

The numerical values are given in Table V for the DC-8 example. Figure 7 is a system survey of the director transfer function numerator (see Eq. 2), consisting of $j\omega$-Bode and Bode root locus plots on the right and a conventional root locus plot on the left. The heavy lines are the Bode which show the variation of closed-loop numerator's real roots with gain, $K_{\dot{h}}/K_{\dot{\theta}}$. The

TABLE V

SELECTED EQUALIZATION VALUES

EQUALIZATION	EXPRESSION	DESIRED LOCATION AND VEHICLE VALUE	SELECTED EQUALIZATION	REMARKS
Pitch Washout	$\frac{1}{T_{WO}}$	$< \omega_{sp} = 1.23$ $> \frac{1}{T_{\theta_2}} = 0.65$	0.7	Washout less than ω_{sp} to provide attitude stability but greater than $1/T_{\theta_2}$ for windproofing and to maintain altitude bandwidth.
Pitch Attitude	$\frac{1}{T_{WO}} + \frac{K_\theta}{K_{\dot{\theta}}}$	$\simeq \omega_{sp} = 1.23$	1.7	Pitch attitude lead set to improve short period damping and extend the region of K/s by having the resulting ω_θ zeros cancel the ω_{sp} poles.
Altitude	$\frac{K_h}{K_{\dot{h}}}$	$\doteq \omega_p = 0.167$	0.2	Greater than ω_p to avoid a "busy" display and the low frequency closed-loop $\dot{d}/d_c$ amplitude droop, yet maintain mid-frequency phase margin.

dotted line along the Bode asymptote is the locus of the complex pair ω_θ. The numerator roots are determined by the gain $K_{\dot{h}}/K_{\dot{\theta}}$. Selecting a value of +0.011 places the complex pair of roots, ω_θ, near the vehicle short-period frequency to cancel ω_{sp} and extend the K/s region.

The open-loop director/vehicle, FD/δ_e, system survey is given in Fig. 8. The equalization terms ($1/T_1$ and ω_θ) do provide the desired K/s-like amplitude ratio over a large frequency region where the pilot should close the loop (the potential crossover region). The actual equalization terms (transfer function zeros) are compared with the approximate characteristic ratios in Table VI.

PILOT LOOP CLOSURE CONSIDERATIONS

In closing the loop, the pilot will introduce a time delay, and perhaps some offsetting high frequency lead. This will modify the open-loop system response properties as shown for an assumed pilot time delay, τ, of 0.4 sec by the dashed phase curve in Fig. 8. This gives the "maximum possible crossover" line which intersects the amplitude ratio plot at about 4 rad/sec. The potential crossover region is sketched in Fig. 8 indicating the extensive frequency band over which K/s-like controlled-element dynamics exist. Nominal pilot-vehicle system crossover will be near the middle of this band at about 1 rad/sec. The actual crossover will vary depending on the pilot, to

$$G(s) = \frac{G_h^{FD} N_{\delta_e}^{h}}{G_\theta^{FD} N_{\delta_e}^{\theta}} = \frac{K_{\dot{h}}/K_{\dot{\theta}}\,(-10.1)(s+.042)(s+.2)(s+.7)(s-3.6)(s+4.4)}{s^2(s+.1)(s+.65)(s+1.7)}$$

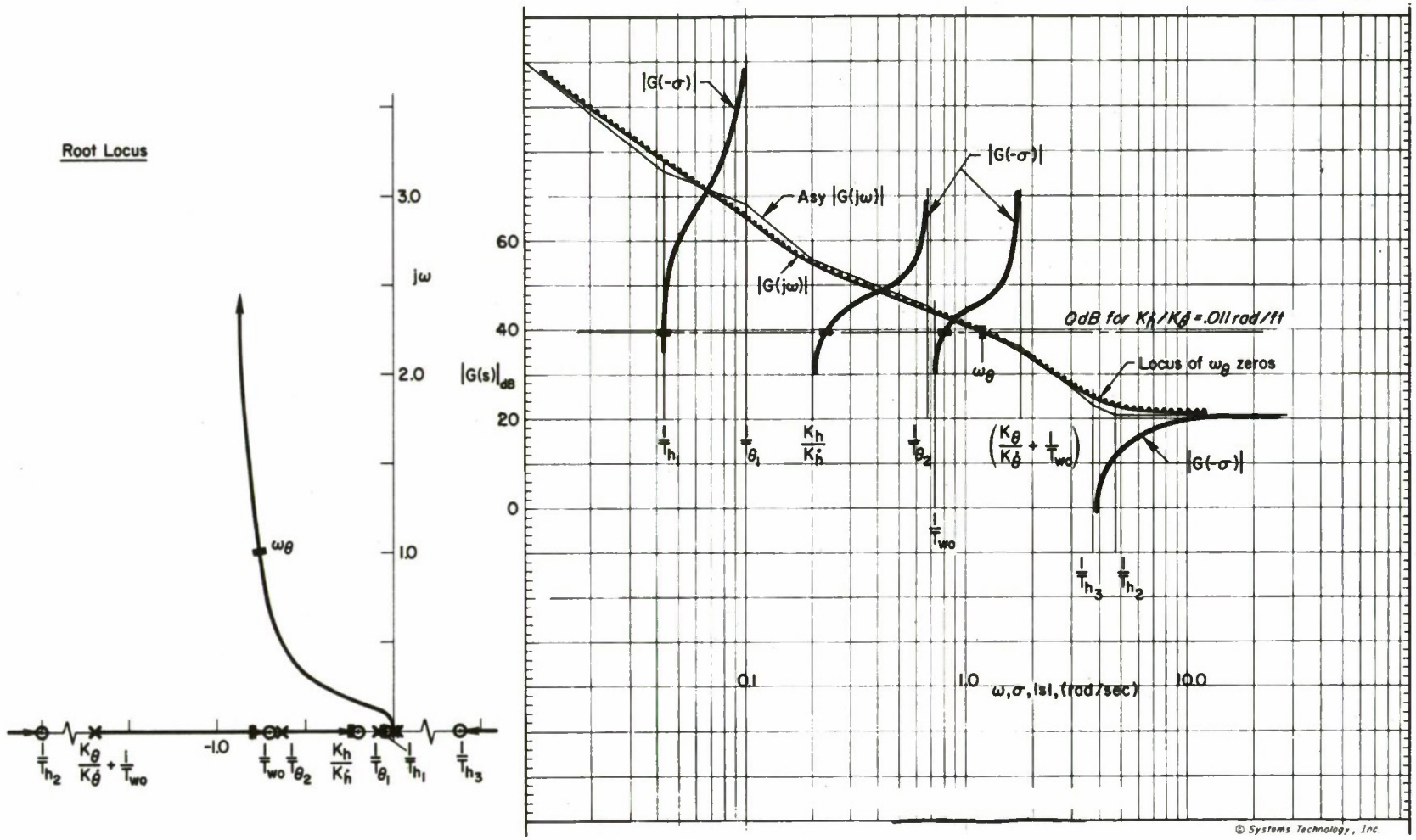


Figure 7. Numerator Survey for Advanced Director/Vehicle Transfer Function

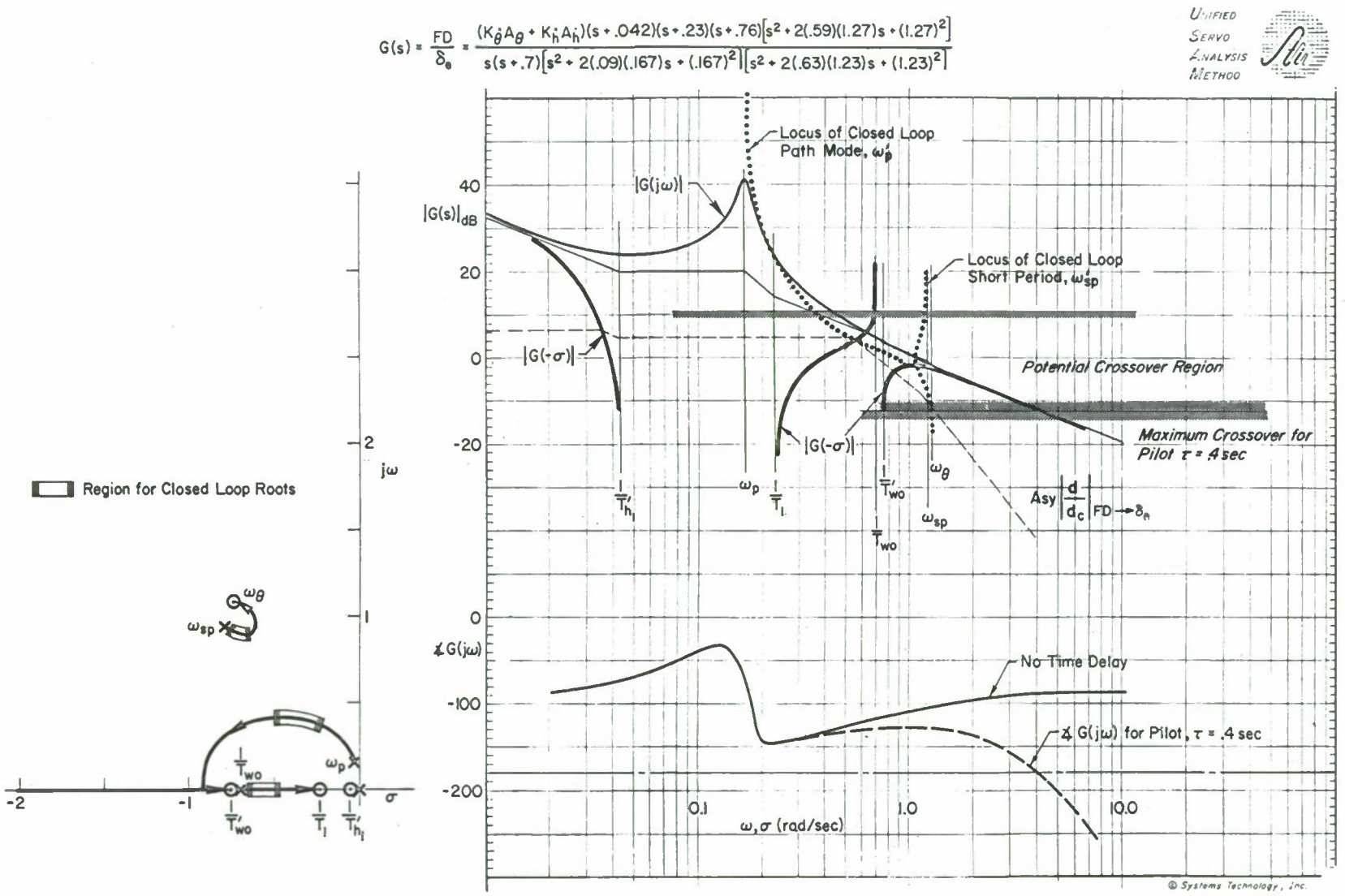


Figure 8. Advanced Pilot/Director/Vehicle System

TABLE VI

COMPARISON OF APPROXIMATE AND EXACT EQUALIZATION ZEROS

EQUALIZATION TERM	CHARACTERISTIC RATIO APPROXIMATION	CALCULATED VALUE
$\frac{1}{T_1}$	$\frac{K_h}{K_{\dot{h}}} = 0.20$	0.23
ω_θ	$\sqrt{-Z_\alpha \frac{K_{\dot{h}}}{K_{\dot{\theta}}}\left(\frac{1}{T_{WO}} + \frac{K_\theta}{K_{\dot{\theta}}}\right)} = \sqrt{(1.89)(1.7)} = 1.78$	1.27

satisfy the guidance and control requirements for particular inputs while at the same time maintaining an acceptable level of vehicle motions (pitch attitude, load factor, etc.). The gain will vary also for different levels of system input (beam bends, turbulence) due to threshold effects. These are more detailed bases for the requirement to provide a broad range of K/s-like dynamics. Then the nature (mode shape) of the system response is insensitive to variations in pilot gain, whatever the cause while the time characteristics change directly with gain.

For large discrete inputs such as an initial beam (step) offset, the pilot's response will differ from that for continuous random inputs. He will tend to operate so that the system responds more rapidly yet with less overshoot (Ref. 11) than the describing function would predict. In the limit, a skilled pilot performing a practiced maneuver may approach a time optimal response, consisting of one well timed and sized elevator pulse in the case of K/s director/vehicle dynamics. When discrete inputs are not dominant, but are mixed with random inputs in a relatively unpredictable way, then the describing function models are appropriate.

For director control of transport aircraft in landing approach, a primary consideration in estimating pilot gain is the frequency and damping of the resulting closed-loop modes. The pilot will be sensitive to the pitch attitude motions required in generating altitude rate motions needed to minimize beam error. One way to describe these motion harmony characteristics quantitatively is with the use of modal response ratios. Imagine, for instance, that the aircraft has been displaced from the beam and that the flight director system is operating to reduce this departure, then the Laplace transforms of the beam response will be given by

$$d(s) = \frac{d_o}{s} + \frac{d_1}{s - s_1} + \frac{d_2}{s - s_2} + \cdots + \frac{d_N}{s - s_N} \qquad (14)$$

where the s_i's are the roots of the closed-loop characteristic equation of the flight director system and the beam forcing function. Then the Laplace transform of other aircraft motion quantities such as attitude, θ, or normal acceleration, a_z, will be

$$\theta(s) = \frac{d_o}{s}\left[\frac{\theta(s)}{d(s)}\right]_{s=0} + \frac{d_1}{s-s_1}\left[\frac{\theta(s)}{d(s)}\right]_{s_1} + \frac{d_2}{s-s_2}\left[\frac{\theta(s)}{d(s)}\right]_{s_2} + \cdots + \frac{d_N}{s-s_N}\left[\frac{\theta(s)}{d(s)}\right]_{s_A}$$

$$a_z(s) = \frac{d_o}{s}\left[\frac{a_z}{d}\right]_{s=0} + \frac{d_1}{s-s_1}\left[\frac{a_z}{d}\right]_{s_1} + \cdots + \frac{d_N}{s-s_N}\left[\frac{a_z}{d}\right]_{s_N} \qquad (15)$$

The bracketed quantities in Eq. 15 are modal response ratios. In general they have both an amplitude ratio and a phase. The closed loop response in a well designed flight director system will be dominated by only a very few (3 or less) basic modes. These will be associated with the system crossover region. The values of s_i within that region are measures of the system bandwidth.

To carry the example further, consider that the crossover frequency is at a location where the short-period equations of motion are approximately valid and that the modal response ratio relating flight path and attitude is pertinent. Under these conditions θ/γ would be given by

$$\left|\frac{\theta}{\gamma}\right| \doteq |T_{\theta 2}s + 1|_{s_i}$$

If the dominant modes, s_i, are such that $|T_{\theta 2}s_i| << 1$, then attitude and path are related on a proportional basis. If the system crossover frequency is considerably higher, $|T_{\theta 2}s_i| >> 1$, then the attitude will be proportional to path rate in this mode.

The system of Fig. 8 presents relatively good vehicle characteristics. With lower short-period damping the potential crossover frequency will have to be reduced in order to stay below any peak in the amplitude ratio. In that case more attitude rate feedback will be necessary. In the opposite sense if the phugoid had had better damping it would not be necessary to use as much K_h gain, i.e., $1/T_1$ could be increased. These effects are reflected in the modal response ratios, along with the effect of varying pilot gain.

Preferred values for the closed-loop modes can be illustrated with the loci of modal response ratios plotted in Fig. 9. The loci of modal response ratios plotted are:

- $|\theta/\gamma|$, pitch attitude to path angle; equal to unity
- $|a_z/\theta|$, normal acceleration at the pilot relative to pitch attitude; equal to 0.1g per degree.

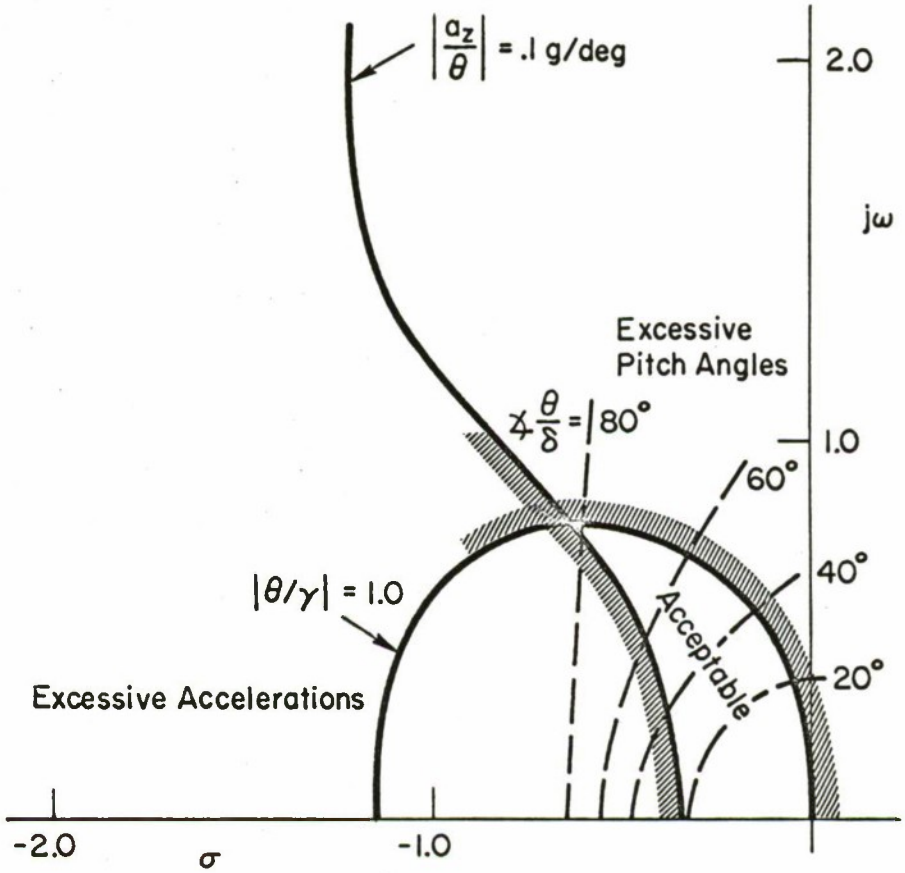


Figure 9. Modal Response Ratio Boundaries for the DC-8 Example

The plotted loci were obtained by evaluating the appropriate open-loop numerator ratios at possible closed-loop modes, and then plotting the frequency and damping associated with a given amplitude ratio (e.g., $|\theta/\gamma| = 1.0$). The loci form hypothetical boundaries for excessive pitch attitude and normal acceleration. Superimposing the example boundaries in Fig. 9 onto the root locus portion of Fig. 8 indicates the allowable gain region for the closed-loop modes. A crossover frequency of 0.6 rad/sec or less keeps the path mode accelerations due to attitude changes less than 0.1g/deg, and the attitude to flight path angle change near unity at short-period frequencies. In other words, the roots on the complex locus in Fig. 8 lie within the example "acceptable" region in Fig. 9.

Another view of the closed-loop motion harmony for this example can be obtained from the beam deviation and pitch angle responses to beam commands. The frequency response plots for a 0.6 rad/sec crossover are shown in Fig. 10. The beam deviation response is flat out to the dominant mode and then rolls off sharply. The attitude response peaks up near the path mode and then rolls off sharply, indicating little attitude overshoot to an altitude command.

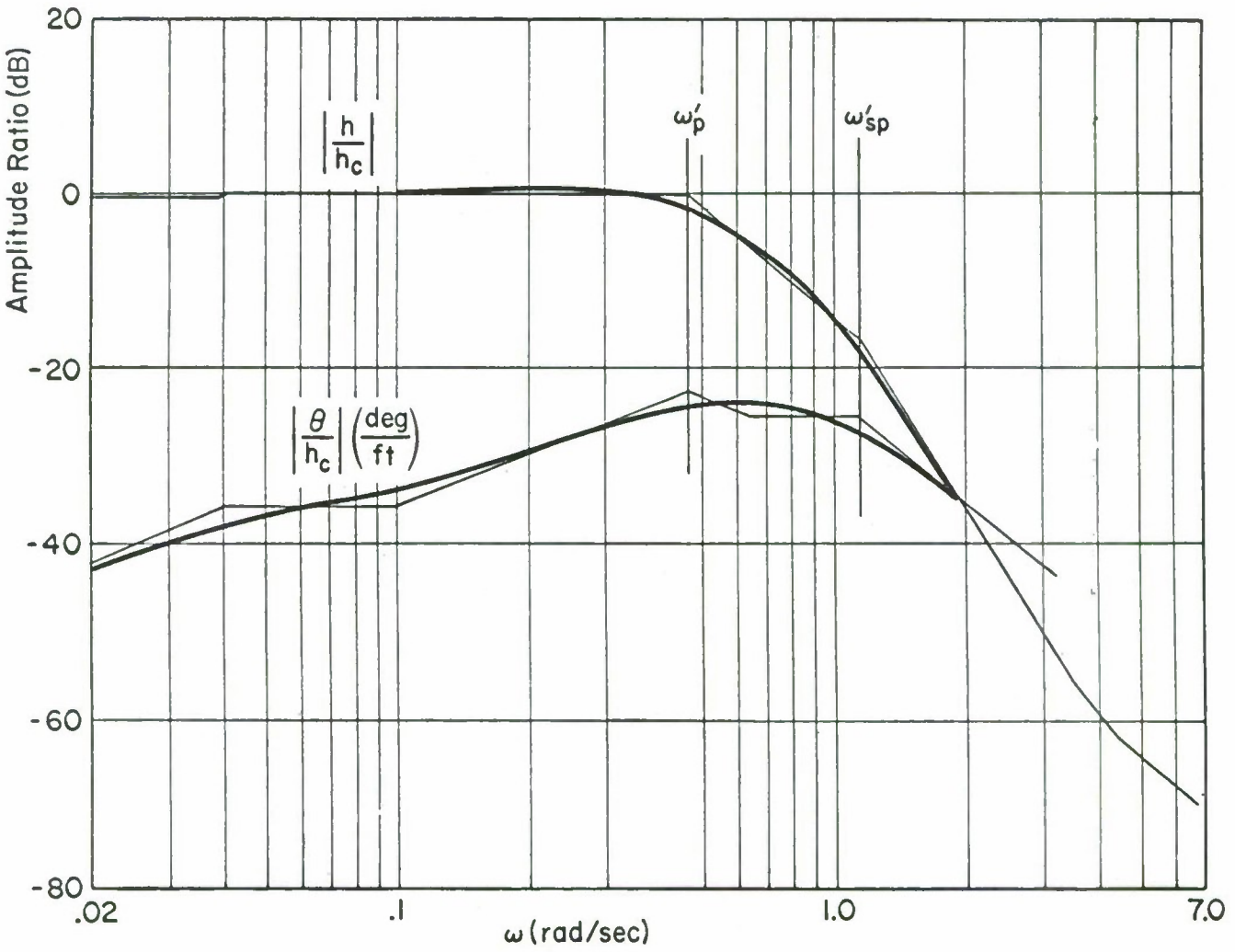


Figure 10. Closed-Loop Attitude and Altitude Response to Altitude Commands for Example Crossover

Other modal response ratios for this example are summarized in Table VII. The θ/h ratio can be used to express the attitude overshoot to an altitude command, by multiplying the θ/h ratio by the h/h_c response at the same mode. The small u/θ ratio at short period indicates that the vehicle holds speed well for short time intervals. At the path mode the speed changes will present little problem to the pilot.

In addition to the path and altitude modes, there will be a closed-loop root at low frequency as the free s at the origin is driven to $1/T_{h_1}'$ (see $|G(-\sigma)|$ on Fig. 8). Although this root is nearly canceled by the $1/T_{h_1}'$ zero in the closed-loop beam response there will be a significant response in airspeed. The airspeed settling time is nearly equal to $1/T_{h_1}$ and increases

TABLE VII

MODAL RESPONSE RATIOS FOR EXAMPLE PILOT CLOSURE

MODAL RESPONSE RATIO	CLOSED-LOOP PATH MODE [0.604 ; 0.467]	CLOSED-LOOP SHORT PERIOD [0.64 ; 1.16]
$\frac{\theta}{\gamma}$ (deg/deg)	$0.65 \angle 47^\circ$	$1.15 \angle 90^\circ$
$\frac{\theta}{h}$ (deg/ft)	$0.06 \angle 162^\circ$	$0.32 \angle 225^\circ$
$\frac{u}{\theta}$ (ft/sec/deg)	$1.75 \angle 19^\circ$	$0.45 \angle -40^\circ$
$\frac{a_z}{\theta}$ (g/deg)	$0.075 \angle 275^\circ$	$0.09 \angle 200^\circ$

as the aircraft approaches the speed for minimum drag unless speed is separately controlled with the throttle.

In summary, these modal response considerations show that the pilot gain and closed-loop properties can be estimated on the basis of what the pilot will consider to be an acceptable repertoire of system responses. This is somewhat different from the usual situation in which the analyst is attempting to estimate pilot gain and system stability margins largely on the basis of predicted path mode error. Unfortunately there are no hard data on the key modal response ratios for flight director or automatic landing systems. Presumably, the value of $|\theta/\gamma|_{s_i}$ for the dominant modes should be near unity to avoid overrotation in corrective maneuvers. As a practical matter, this is less of a problem with flight director systems than with automatic approach systems because the pilot is free to select the gain he feels is suitable in closing the loop.

CONCLUSIONS

A set of principles, functional requirements, and analytical procedures can be defined for specifying and designing flight director/vehicle systems. These permit the designer to select, equalize, and weight the director feedbacks analytically, given the (augmented) vehicle dynamics and a definition of the task. This involves some tradeoff between the basic guidance and control requirements and additional requirements based on human pilot considerations. These new concepts include the following:

- The effective director/vehicle controlled element should look like a K/s over a broad mid-frequency region.

- The director display should be consistent with status information — low frequency and steady-state bar motions should be beam deviation, the mid-frequency deviations should reflect corresponding vehicle motions, and high frequency motions should be attenuated.

- The harmonization of attitude and path motions has an important influence on pilot gain and system crossover frequency.

- Scanning required to monitor status information will tend to reduce pilot gain and this can be avoided by suitably integrating the status information on the display.

By using these design principles and analytical techniques the final optimization process using actual pilots during simulation and flight test can be planned and accomplished more expeditiously.

REFERENCES

1. McRuer, Duane, Henry R. Jex, Warren F. Clement, and Dunstan Graham, A Systems Analysis Theory for Displays in Manual Control, Systems Technology, Inc., Tech. Rept. 163-1, Oct. 1967 (revised June 1968).

2. Clement, Warren F., Henry R. Jex, and Dunstan Graham, "A Manual Control-Display Theory Applied to Instrument Landings of a Jet Transport," IEEE Trans., Vol. MMS-9, No. 4, Dec. 1968, pp. 93-110.

3. 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., Tech. Rept. 170-2, July 1969 (forthcoming NASA CR-).

4. Weir, David H., and Richard H. Klein, The Measurement and Analysis of Pilot Scanning and Control Behavior During Simulated Instrument Approaches, Systems Technology, Inc., Tech. Rept. 170-4, June 1969 (forthcoming NASA CR-).

5. Clement, Warren, and Lee Gregor Hofmann, A Systems Analysis of Manual Control Techniques and Display Arrangements for Instrument Landing Approaches in Helicopters. Vol. I, Speed and Height Regulation, Systems Technology, Inc., Tech. Rept. 183-1, July 1969.

6. Johnson, Walter A., and Duane T. McRuer, Development of a Category II Approach System Model, Systems Technology, Inc., Tech. Rept. 182-1, Dec. 1969.

7. Weir, D. H., R. H. Klein, and D. T. McRuer, Principles for the Design of Advanced Flight Director Systems Based on the Theory of Manual Control Displays, Systems Technology, Inc., Tech. Rept. 170-5, Mar. 1970.

8. 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.

9. 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.

10. McRuer, D., and D. H. Weir, "Theory of Manual Vehicular Control," Ergonomics, Vol. 12, No. 4, 1969, pp. 599-633.

11. McRuer, D. T., L. G. Hofmann, H. R. Jex, G. P. Moore, A. V. Phatak, D. H. Weir, and J. Wolkovitch, New Approaches to Human-Pilot/Vehicle Dynamic Analysis, AFFDL-TR-67-150, Feb. 1968.

A STUDY OF RELATIONSHIPS BETWEEN AIRCRAFT SYSTEM PERFORMANCE AND PILOT RATINGS

William C. Schultz, Frederick D. Newell and Richard F. Whitbeck
Cornell Aeronautical Laboratory, Inc.
Buffalo, New York 14221

ABSTRACT

This study examines the relationship between man-machine system performance and pilot evaluation data. The intent of the experiment described herein is to add to the library of knowledge concerning quantitative analytical measures of pilot-aircraft performance for complex piloting tasks.

INTRODUCTION

In the course of a previous study, described in Reference 1, a simulation experiment was performed which employed analytic performance measures in an attempt to obtain objective measures of overall system performance. In that experiment, critical parameters in the aircraft equations of motion were varied and subjective pilot evaluation data was taken. After computing a non-parametric Spearman rank correlation coefficient between certain man-machine performance values and pilot evaluation data, it was discovered that the subjective and objective performance measures appeared to be highly correlated. Very high correlation values were obtained, for example, for a measure of pitch error, or pitch error and stick motion. Thus, in the relatively simple experiment that was performed, there appeared to be a strong correspondence between man-machine performance and pilot rating.

The investigation* reported in this paper is an expansion of the earlier effort, conducted to study in greater depth, objective and subjective pilot performance measures. Specifically, would the hypothesis developed on the basis of the aforementioned, relatively simple longitudinal experiment (i.e., a high correlation between an analytic measure of performance and the equivalent pilot rating) continue to hold when a more complex, combined longitudinal-lateral directional ILS approach task was attempted.

In this more complex experiment, a fixed base ground simulator was used. Care was taken to obtain a realistic simulation by using real cockpit displays, comprehensive equations of motion and experienced evaluation pilots. The basic approach was to obtain pilot rating data in sufficient quantities to ensure that the statistics could be thoroughly examined. Then hypothetical performance indices, similar to the ones used in the original (simpler) experiment, were chosen and the same statistics re-examined. The two sets of data can be considered to be statistically sensitive to the same degree, if the performance index data displays the same statistics as the pilot rating data.

---

*This work was sponsored by Langley Research Center, NASA, under Contract No. NAS1-8765.

The results of this more complex ILS approach experiment were completely negative for the limited set of performance measures that were examined. That is, the high correlation that was suggested on the basis of the simpler longitudinal tracking task failed to materialize for the ILS approach task. A brief description of this ILS landing task and the results obtained, are given in the sections which follow.

APPROACH USED IN THE EXPERIMENT

The task was to fly an ILS approach from outer marker to middle marker. To properly load the pilot, both longitudinal and lateral-directional dynamics were used as well as a full instrument presentation and operable throttle, elevator and elevator trim, aileron, and rudder controls. Side gusts were provided in the lateral-directional modes through sideslip terms in the equations of motion. Although gusts were not injected into the longitudinal equations, the activity and concentration required of the pilot to maintain heading caused him to introduce noise into the longitudinal motions, mainly by his having to realistically divide his attention among the many flying sub-tasks.

In evaluating each configuration, the pilots rated the longitudinal mode only, the lateral-directional mode only, the total overall airplane, and whether or not the airplane could be landed.

Five pilots performed the experiment. Four pilots performed each set of configurations three times and the fifth pilot performed each set twice. The configurations were performed in a random order; no pilot had the same random order twice, and no pilot had a random order that had been assigned another pilot. Each pilot was given a training period of one to two hours before data was taken. Then, for every data run with a specific configuration, there were two preliminary practice runs in which the pilot was free to make maneuvers (including getting off glide slope, and/or making large corrections) to help him determine the characteristics of the configuration.

The only task that was rated was the ability to maintain the glideslope and localizer acceptably well for the ILS task. There was no sensation of "g" and no "g" meter was used in the simulator.

1. Simulation of Aircraft Configurations

A fixed-base aircraft simulator was used in the study. The cockpit instrumentation included: airspeed, altitude, pitch angle, bank angle, heading, glide slope and localizer, outer marker, flap position, turn rate, slideslip ball, rate of climb (VSI), and %RPM. Aircraft dynamics were introduced through a set of six degree-of-freedom equations. An Electronics Associates, Inc. TR-48 analog computer and a COMCOR CI-5000 analog computer were used to simulate the aircraft equations of motion and to drive the aircraft instruments.

Limited six degree-of-freedom equations of motion were selected for the study. These equations provided control through elevator, aileron, rudder and throttle (δ_x) and changes in flap settings (δ_f). Elevator trim was also provided.

The emphasis in the program was on the longitudinal short-period undamped natural frequency and damping ratio combinations. For this reason, the phugoid dynamics were held constant at a damping ratio of 0.05 and a period of 35 seconds.

Lateral-directional characteristics were selected to be representative of a typical large airplane. Again, the emphasis on the longitudinal short-period characteristics made it necessary to keep the lateral-directional characteristics constant.

A constant stick force per "g" was used for each configuration and then, for a smaller set of configurations, it was varied naturally with $\omega_{n_{sp}}^2$. The stick spring rate (F_{es}/δ_{es}) was constant.

The short-period characteristics evaluated were as follows:

ζ_{sp} \ $\omega_{n_{sp}}$	0.5 rad/sec	1.0	3.0	6.0
0	✓	✓	✓	✓
.3			✓	✓
.6	✓	✗	✓	✗
1.0		✓	✓	✓

The twelve conditions (✓ or ✗) were done with the same F_{es}/g chosen to be typical for the $\omega_{n_{sp}} = 1.0$, $\zeta_{sp} = 0.60$ case. The two conditions marked (✗) were run holding M_{δ_e} constant and allowing F_{es}/n_z to vary with M_α. Again, the base condition was for $\omega_{n_{sp}} = 1.0$ and $\zeta = 0.60$.

2. The Pilot Task

The task was an ILS approach to landing in light to moderate turbulence. The pilot was given the airplane at the glide slope intercept altitude of 1500 feet, approximately one mile out from the outer marker. The airplane was trimmed and on the localizer. The task was to fly to the outer-marker (glide slope intercept) at 1500 feet and then to follow the localizer and glide slope down to 200 feet altitude. A flare was not made. All pilots were currently active test pilots experienced in handling qualities evaluations.

The pilot was told to fly a maximum-safety, minimum-deviation ILS approach as he would do in a large airplane. He was told to take the airplane down to fifty feet, although data was used down to only 200 feet. This assured that the pilot was paying maximum attention to the task through the end of each data run. The minimum deviation aspect of the approach was also stressed, because not all pilots actually fly a minimum deviation approach. Some pilots fly a path more like the one a flight director would compute. By stressing the minimum deviation criteria, each pilot would tend to fly by the same rule.

3. Pilot Comment Card and Rating Scale

The pilots' understanding of the specific task to be done and the purposes of the comment card are crucial to obtaining consistent and interpretable pilot evaluation data.

A card that educes sufficient consistency in the comments without overconstraining the pilot is desirable. If the comment card is too constraining, then data will be lost. Yet the ILS task is a maximum attention task and the pilot cannot recall, after the task, all of the subtleties he noticed during performance of it. To record their comments, the pilots were given a live mike while they performed the task, yet the concentration on the task was so high that none of them were able to comment while performing the task. In fact, they found that to do so affected task performance. Therefore, pilot comments were obtained immediately at the conclusion of each approach.

Another purpose of the comment card was to determine how the pilot reacted to the individual variables he perceived, and to the workload required to accomplish the task. The comment card used is shown in Figure 1. The rating scale used was the Cooper- Harper, ten point scale of Reference 2.

Task - Large Transport, Minimum Error ILS
Pilot Comments

Pilot _____ Run No. _____ Date _____

Comment On:

- Forces
- Pitch Response
- Control of Pitch Attitude
- Control of Pitch Rate
- Altitude Control
- Altitude Rate Control
- Control of Glide Slope - Glide Slope Rate
- Throttle Technique

Lat.-Dir. Comments:

Rate:

1. Longitudinal Only
2. Lat.-Dir. Only
3. Overall Airplane
4. Is Airplane Landable?

Figure 1. Pilot Comment Card

EXPERIMENTAL DATA OBTAINED

Data were taken for all five pilots for each of the fourteen configurations used. These data exist in three forms: analog records on two ten-channel Brush Recorders; digital tape recordings were made on 22 channels; pilot comments were recorded after each run. Thus, a vast amount of data was recorded, in several forms, during the experiment. These data are described below.

1. Analog Records

One of the ten-channel recorders was used to record eight-channels of data from the longitudinal mode. These included elevator inputs, throttle inputs, pitch rate, pitch attitude, airspeed, rate of change of altitude, altitude, and glide slope. A second recorder was used to record eight-channels of data from the lateral-directional mode. These included rudder and aileron inputs, sideslip, bank angle, heading and heading rate, vertical acceleration, and the localizer error. A typical set of data is shown in Figure 2.

2. Digital Tape Records

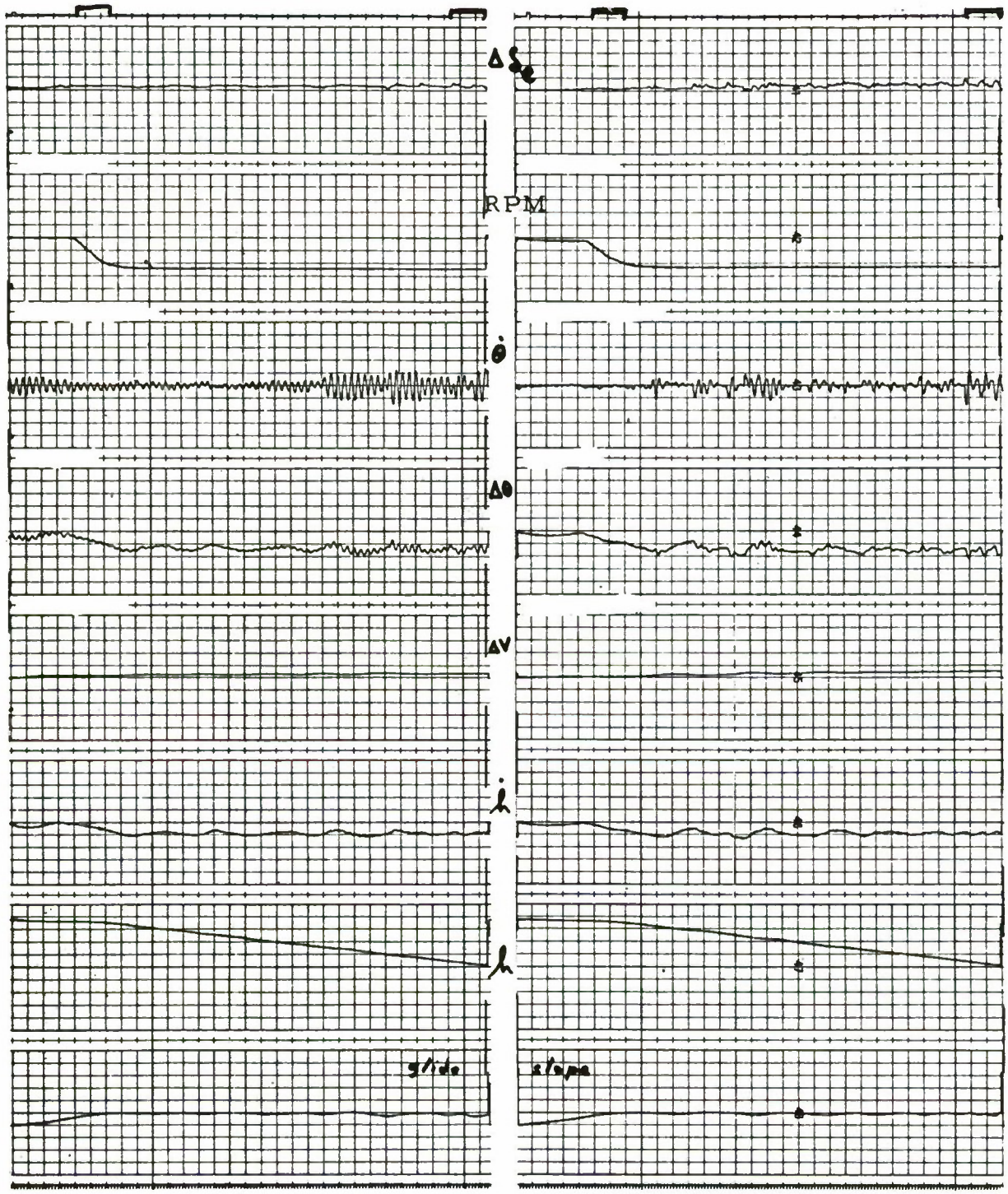
Real-time digital recordings were made for all runs of the experiment. An off-line analog-to-digital converter multiplex digital recording system was used to record twenty-two channels of data. These included all those shown on the analog records plus elevator trim, angle-of-attack, and gust disturbance input.

3. Pilot Comment Tape Records

In this experiment, pilot comments were recorded after each run. These comments proved to be one of the most important outputs of the experiment. The comments explained the various "tricks" and techniques that the pilots employed to obtain small glide slope deviations even though many of the configurations were extremely difficult to handle. For example, in the zero damping, short-period cases the comments revealed that the pilots intentionally did not use the elevator control for fear of exciting the short-period mode. This is considered to be a case of "extreme pilot compensation" when a standard, normal control has to be avoided. In many of these extreme cases, the glide slope performance was comparable to that achieved in cases where normal control techniques were employed.

4. Pilot Rating Data

As mentioned earlier, in evaluating each configuration, the pilots rated the longitudinal mode only; the lateral-directional mode only; the total overall airplane; and whether or not the airplane can be landed. The data is in tabular form. A typical table, for pilots' ratings of the longitudinal mode only, is shown in Table 1.



(a) (b)

Pilot Rating 9 Pilot Rating 2 1/2

FIGURE 2a Longitudinal Data $\zeta_{sp} = 0$; $\omega_{sp} = 3.0$

FIGURE 2b Longitudinal Data $\zeta_{sp} = 0.3$; $\omega_{sp} = 3.0$

FIGURE 2. Typical Set of Analog Records

TABLE 1

Pilot Ratings

Longitudinal Only

Pilot	Set	ζ_{sp} = 0				.3		.6					1.0		
		ω_{sp}=0.5	1	3	6	3	6	1	3	6	*1	*6	1	3	6
1	1	10	10	10	8	4	4	7	3	3	7.5	5.5	7	4	3
	2	9	10	9	9	2.5	3	4	2	3	7	7.5	7	3	3
	3	9.5	10	9	8.5	2.5	2.5	5	2.5	3	9	6	4.5	2	4
2	1	10	8	7	7	4.5	7	5	3	2	7	7	5	4	3
	2	10	10	6	6	5	3	7	2	4	8	9	8	4	2
	3	10	10	6	5	3	2	5	4	3	7	4	4	4	4
3	1	10	9	7	9	4.5	6	5	6	4.5	6	6	6	6	6
	2	10	8	8	7.5	5	6	3	5	4	8	4.5	6	4.5	7
	3	9	8	7.5	8	5.5	5	7	3	4.5	7.5	7	7	4	4
4	1	10	8	7	8	4	6	5	6	4	10	7	6	5	8
	2	10	10	7	7	4	2	8	5	3	10	5	5	4	4
	3	10	10	7	7	3.5	3	4	3	2	10	6	6	5	3
5	1+	10	10	10	10	3.5	5.5	7	5	5.5	10	9.5	9.5	6.5	5
	2	10	10	10	10	3	6	9	5	4	10	9	9	9	6
	3	10	10	10	10	4	5	5	5	7	10	10	10	4	4

* Changed Stick Forces

\+ Pilot 5 performed only two sets of data runs
The data shown here for Set 1 are averages of his Sets 2 and 3

ANALYSIS OF DATA AND RESULTS

The data format of this experiment is directly amenable to use of the analysis of variance. This analysis technique is used to determine the statistics of the pilot rating data. The entire format of data is not orthogonal as it stands, because not each of the frequencies was examined for every damping ratio that was used. Therefore, two different sub-sets of the data, which give orthogonal comparisons, are used in the analysis of variance.

The first sub-set of pilot ratings or objective measures analyzed (Case I) was for the short-period frequencies of 1, 3 and 6 radians/second, each at damping ratios of 0, 0.6 and 1.0. Because each pilot did each set of runs three times, a test for learning may be made. The main factors in the analysis are, therefore, learning, frequency, and damping ratio.

The second sub-set of pilot rating data analyzed (Case II) was for the short-period frequencies of 3 and 6 radians/second each at a damping ratio of 0, 0.3, 0.6, and 1.0.

Pilot Rating Data

In Case I, the analysis shows that learning is not significant, and that both frequency and damping ratio are significant at the 5% level. This sub-set of the data was analyzed in both the raw and normalized form, and the results are the same for both analyses.

For Case II, both in the normalized and raw forms, the results are that both learning and damping ratio are significant at the 5% level and frequency is not significant.

In each of the analyses of variance, none of the interactions between or among set, frequency, or damping ratio are significant. The model for these analyses include the pilots in the error term; and, therefore, there is no measure of significant differences among the pilots.

Analysis of Glide Slope rms Data

The analysis of variance of Case I, using rms glide slope error, shows results which agree with the similar analysis of the pilot rating data for the main effects of frequency and damping. However, it also adds the interaction between frequency and damping ratio as significant at the 5% level.

For Case II, the sub-set of 3 and 6 radians, for four damping ratios, the results are very much different. In this case there were no significant main effects or interactions.

It is interesting that the Case I glide slope rms data shows significant trends, especially since the glide slope rms does not correlate with pilot rating as shown in Figure 3. An investigation of the data shows a strong interaction for the zero damping ratio, one radian/second point and it is this single point that gives rise to the significance shown. The significance can, therefore, be considered spurious.

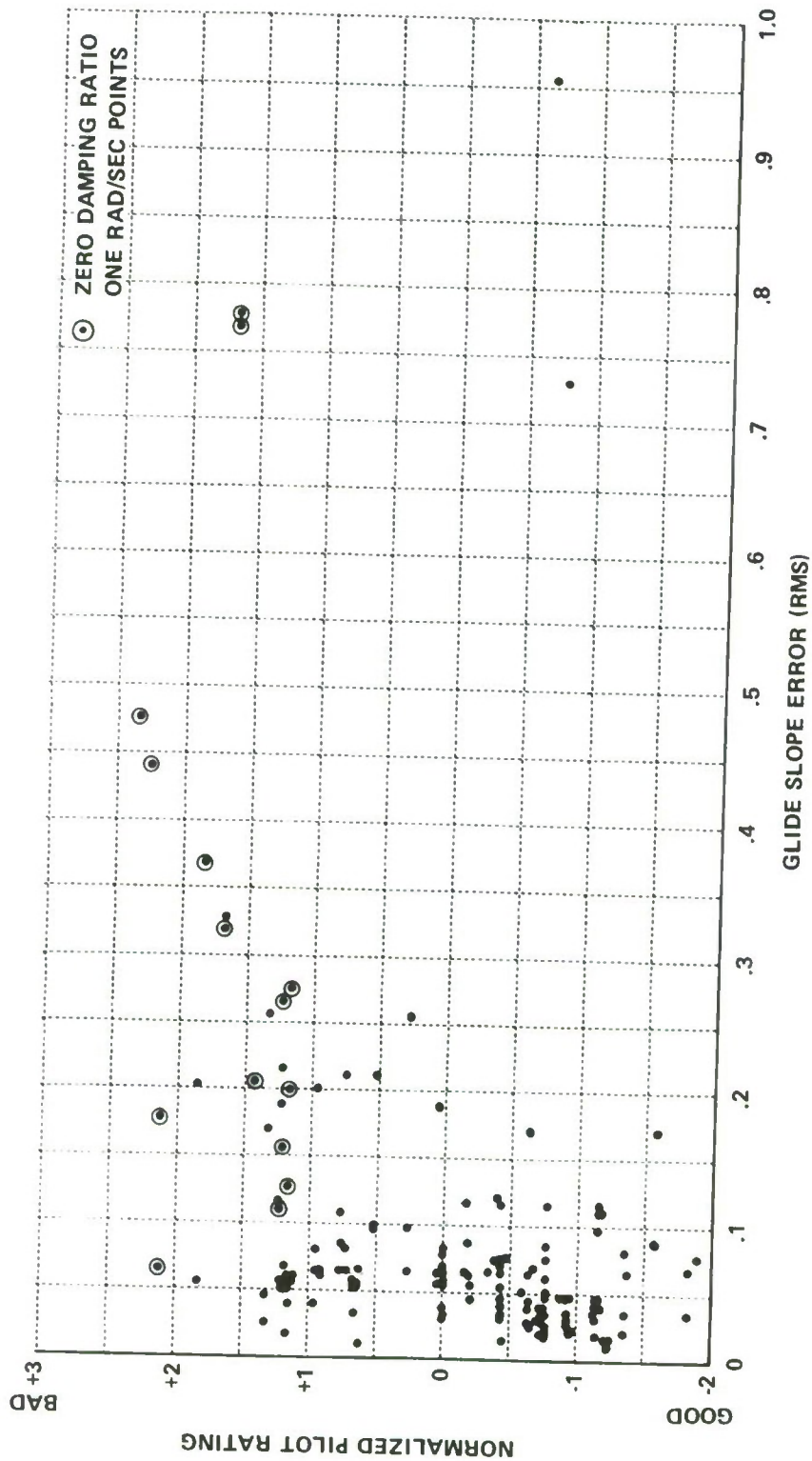


Figure 3 PILOT RATING VS. GLIDE SLOPE ERROR

Pilot Comments

The use of a "live mike" was attempted to record pilot comments during the data runs. It was conclusively established that attempting to comment during the run degraded their performance. Hence, after a few attempts, the pilots refrained completely from making comments during the run.

From the pilots' tapes, it has been established that the pilot rating is dependent on distance inbound from outer marker. For example,

a. They frequently rate the configuration as poorer because of relatively poorer performance as they approach the middle marker.

b. Pilot rating is not readily apparent from records of glide slope error. The pilots fly a poorer configuration more tightly than a good one because they are afraid of "losing it." With a relatively well behaved configuration, they will tolerate more error since recovery is easier.

In addition, the pilot comments indicate that the same pilot will employ a variety of flying techniques for the different configurations.

There is no correlation between pilot rating and rms glide slope error.

REFERENCES

1. Wierwille, W. W. and Knight, J. R. Synthesis Methods for Manual Aerospace Control Systems With Applications to SST Design. CAL Report No. IM-2429-B-1, Project WEIWOHO, Contract No. NAS1-7141, March 1968.

2. Cooper, G. E. and Harper, R. P. The Use of Pilot Rating in the Evaluation of Aircraft Handling Qualities. NASA TN D-5153, April 1969.

"Paper Pilot" -- An Application of Pilot Models to Predict VTOL Flying Qualities in Precision Hover

R. O. Anderson and J. D. Dillow
Air Force Flight Dynamics Laboratory

ABSTRACT

A study of the correlation of measured pilot model parameters and closed-loop performance with pilot opinion of VTOL hover dynamics was conducted and a method of calculating pilot ratings from the data was developed. This result was implemented to establish a methodology for predicting pilot model parameters, closed-loop performance, and pilot ratings for the hover task. This methodology is called the minimum pilot rating analysis method since it is based on the hypothesis that the pilot adjusts his method of control to minimize pilot rating. A digital computer program, dubbed the "paper pilot", was developed to automate the minimum pilot rating analysis method.

The "paper pilot" rating was computed for 24 aircraft configuration/gust intensity combinations. These ratings are compared to actual pilot ratings obtained in fixed base simulation. The difference between the actual pilot ratings and the "paper pilot" ratings has a mean of .12 and a standard deviation of .55 on a 10 point Cooper Rating Scale.

INTRODUCTION

With an ever increasing reliance on flight control augmentation to provide acceptable flying qualities for high performance, multi-purpose design military aircraft, two problems arise in respect to the specification of flying qualities; namely,

a. To treat the specification problem as an extrapolation of past experience, a whole array of possible types of augmentation must be considered, with attendant requirements for each. Not only are data difficult to amass, but the lack of an underlying set of "true" requirements also becomes apparent. That is, one should certainly state what is really required, and this should apply to any and all means of obtaining these "basic" requirements.

b. On the other hand, the current conventional vehicle specifications are based on the a priori assumption that the aircraft flies "like an airplane". If the vehicle dynamics are indeed augmented, the resulting "effective" vehicle dynamics could easily present seven or more aperiodic modes in place of the conventional longitudinal pair of oscillatory modes. The former may even yield better pilot acceptance, but the question remains, "Does the effective system meet the specification?"

Although both of the above problems have become increasingly more acute over the past few years, the "normally augmented" VTOL hover task has brought things to a head. In short, some new approach must be evolved.

As part of a joint AF-Navy-Army effort to draft a V/STOL flying qualities specification, a new approach to the specification of hover flying qualities was evolved and tested. This evolution is discussed in the remainder of this paper, along with a new pilot-vehicle analysis method that has implications far beyond the specification of VTOL hover dynamics.

PILOT RATING AS A FUNCTION OF PILOT MODEL PARAMETERS AND CLOSED-LOOP PERFORMANCE

The concept of mathematical models of the human operator, or pilot, has developed over the years to play a prominent role in modern manual control theory. Furthermore, the suggestion that pilot model parameters are, in turn, related to pilot opinion rating is not new (Reference 1). For the hover task, Reference 2 discusses the relation in general, while References 3 and 4 indicate specific trends. It remains, however, to consolidate these ideas into a complete correlation, and Reference 5 provides further data for this purpose.

It has been generally observed (Reference 5) that pilot opinion rating is directly related to pilot lead generation (T_L) with the trend of an increased numerical rating (less desireable) with with increase in lead. Also, a trend of increased rating with an increase in closed-loop mean absolute error is evident in Reference 5. Although the piloting tasks considered in Reference 5 did not include VTOL hover under gusty conditions, the pilot rating data in References 3 and 4 show similar trends. Also, the latter references include measured pilot model parameters for an assumed pilot form that produced a good fit to measured values of closed-loop pilot-vehicle performance.

Assuming that lead generation and closed-loop performance are the main factors (there are certainly many others) that affect pilot rating of the longitudinal dynamics in hover under the influence of random gust inputs, the data in Reference 3 were used to develop the following expression for pilot rating as a function of the above factors:

$$R = R_1 + R_2 + R_3 + 1.0$$

where:

$$R_1 = \frac{\sigma - \sigma_m}{\sigma_m}, \quad \sigma_m = 0.80 = \text{constant}$$

$$\sigma = \sigma_x + 10\sigma_q$$

$$0 \leq R_1 \leq 2.50$$

and σ_x = standard deviation of x displacement in feet

σ_q = standard deviation of pitch rate in rad/sec.

also, $R_2 = 2.5\, T_{L_\theta}$

$R_2 \leq 3.25$

T_{L_θ} = Pilot lead time constant in pitch (seconds)

and $R_3 = 1.0\ T_{L_x}$

$R_3 \leq 1.20$

T_{L_x} = Pilot lead time constant in displacement (seconds)

Pilot rating is a numerical value from 0 (best) to 10 (worst) based on the Cooper rating system (Reference 3). Pilot ratings, PR_H, for the cases where measured pilot parameters exist, are shown in Figure 1 versus R, the rating calculated with the above expression, for the data in Reference 3.

Although the rating expression was obtained as an "eyeball fit" to the data (instead of using a more elegant linear regression analysis) the results in Figure 1 show a correlation within one-half rating unit in most cases. In three cases the correlation difference is greater than one rating unit, but:

a. one pilot rated a configuration within 0.12 rating units of R, while the other pilot rating, indicated by a line connecting the two points, differed from R by more than one unit.

b. in two cases large simulated gust intensities seemed to affect the fit. The standard deviation of the horizontal gust, σ_{u_g}, was 5.1 ft/sec except for those cases noted by the value in parentheses next to the data points in Figure 1.

In any event, the majority of data correlate with the rating expression within less than one rating unit, which is within the accuracy of most flying quality analyses. It should also be noted that a rating range of about 2.0 to 6.5 is included (Figure 1).

Based upon the earlier discussion in this Section it would be nice to state that "pure theory" yielded the R expression. Actually it was a simple "curve fit" to the data. Only the form (i.e., R_1 reflects performance, R_2 and R_3 reflect pilot "work load") was "theoretical", as indicated above. The idea of increased rating with increased lead generation (R_2 and R_3) is not new for this task, and others (Reference 2). However, numerical values are difficult to come by. The second form, R_1, is also not new since various performance measures have been used in attempted correlations with pilot rating, "work load", etc. Perhaps what is new here, at least to the authors, is that both performance <u>and</u> pilot leads are used <u>quantitatively</u>. Nonetheless, other analyses (Reference 2) <u>may</u> have used some indirect measure of performance, and Reference 6 does relate closed-loop time constants to pilot rating. Gust effects are not treated, however. At any rate, each term is discussed below, for the most part, in hind-sight:

R_1: $2.5 \geq R_1 = \dfrac{\sigma - \sigma_m}{\sigma_m} \geq 0$. This expression is an attempt to incorporate several "logical" factors. First, some minimum performance, σ_m, is necessary to do a given task. If the actual closed-loop performance, $\sigma = \sigma_x + 10\sigma_q$, is equal to, or smaller than σ_m, no increase in pilot rating is indicated. The normalized form shown was selected because σ_m probably varies with the task. The value of σ_m, for the hover task, was taken as the lowest value of $\sigma = \sigma_x + 10\sigma_q$ encountered in the data. It was equal to about 0.70 for one pilot, and 0.90 for the second, so an average of $\sigma_m = 0.80$ was used. The form of σ itself was simply in keeping with the data. The σ_x part is obvious, since the task is to keep x = 0. However, very large excursions in q would represent excessive

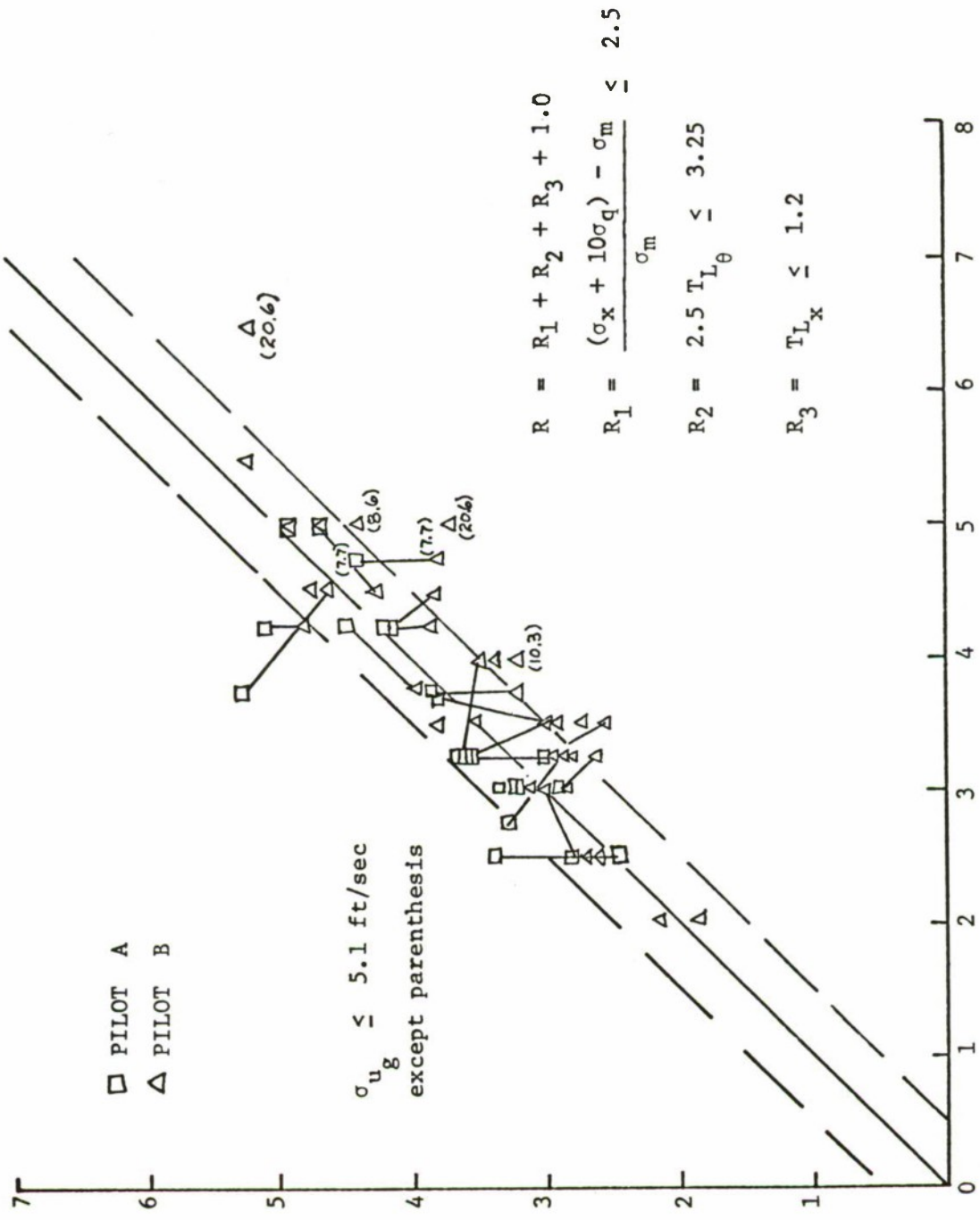


Figure 1. Pilot Rating vs. R

attitude changes occurring during the process of keeping x small, so the $10\sigma_q$ was added as an inner-loop performance measure. The factor 10 just seemed to fit the data and produces rating increments that are twice as large for σ_x as for σ_q with a "fair" configuration.

The unity coefficient of the normalized performance term R_1 does correspond with other data. Figure 27, Reference 5, shows pilot rating vs. mean absolute error, with only the input variance changed (i.e., controlled element fixed). Using $|\overline{e}|$ as an approximation to σ, the data in this figure show the following for the integrator controlled element:

$$R_1 = k\frac{(\sigma - \sigma_m)}{\sigma_m} \qquad \text{or} \qquad 2.5 \doteq k\left(\frac{1.25 - .4}{.4}\right)$$

so k = 1.18 if the $|\overline{e}|$ = .4 is taken as the normalization value.

This indicates the "curve-fit" coefficient of R_1 is, at least, reasonable in comparison with other data. The slope $\Delta R/\Delta|\overline{e}|$ does not correspond very well with $\Delta R_1/\Delta\sigma$, however, so some form of normalization appears necessary for general applicability.

The upper limit on R_1 such that $R_1 \leq 2.50$ is empirical. It fits the hovering data; in particular cases PH 18 - 22 (Reference 3). Two possible, but weak, arguments for this upper limit follow:

a) If no pilot leads are required, then the maximum R determined by performance alone becomes R = 3.5 (Figure 27, Reference 5, shows no apparent limit, however). This happens to be the"acceptable-unacceptable" boundary and, perhaps, beyond this "who cares".

b) If each of the limiting values of R_1, R_2, and R_3, are reached, then $R = 7.95 \doteq 8.0$. All the data used to generate the R expression were ratings for the longitudinal axis with good dynamics in the lateral axis. In turn, the "best" dynamics were rated 2.0. If this 2.0 represented the lateral axis, then the total would be R + 2.0 = 10.0 maximum. Unfortunately, this reasoning leads one to conclude that a 6-axis task is rated 6.0 at best (i.e., a bias of 1.0 per axis).

R_2: $R_2 = 2.5\ (T_{L_\theta}) \leq 3.25$

Fortunately, more effort has been expended on lead term effects than "normalized performance" effects, and the R_2 expression comes directly from Reference 5. If the smallest T_L data point and the next to largest T_L point in Figure 22, Reference 5, are joined by a straight line, this line intercepts the maximum pilot rating increment point at a lead of about 1.3 seconds. The maximum increment in rating is about 3.25 units, so

$$R_2 = k(T_L), \quad \text{or} \quad 3.25 = k(1.3)$$

so k = 2.50.

The use of a straight line approximation to the complete curve is for simplicity.

R_3: $R_3 = 1.0(T_{L_\theta}) \leq 1.20$

At first it might seem that R_3 should be the same as R_2. The data do not support this idea. This may be due to the fact that the pitch loop (with T_{L_θ}) has a high bandwidth (cross-over frequency about 3 rad/sec, Reference 3) for the situation in which R_2 was obtained. However, the outer loop (with T_{L_x}) has a much smaller bandwidth of about 1 rad/sec. Perhaps maximum lead generation induced increments in rating are directly proportional to the bandwidth of the loop in question (in this case 1/3 of the R_2 maximum increment of 3.25 is about 1.10, or very close to the observed T_{L_x} induced increment maximum of about 1.20). On the other hand, considering the full longitudinal-lateral case, four lead terms can be generated. If two of these follow the Reference 5 data for the high frequency loop, and performance is not considered, then maximum pilot rating for the second two leads can be found from the maximum rating of 10.0 as:

$$PR = 10 - 2(3.25) - 1.0$$

$$= 2.50$$

This "leaves" a value of 1.25 as maximum increment for each of the two low bandwidth loops, and <u>any stable</u> performance might be acceptable under these conditions (i.e., $R_1 = 0$).

In any event, the maximum increment in pilot rating is considered to occur at about the same value of lead as in R_2, so the unity coefficient on the R_3 term is used.

Bias: $R = R_1 + R_2 + R_3 + 1.0$

The unity addition to form the final rating expression is simply a result of the rating scales used to collect all of the data (see Reference 3). A rating of 1.0 is described as "excellent, includes optimum", and therefore becomes, for the most part, a lower bound on the numerical rating. In short, with no lead generation and adequate performance, R = 1.0.

For more details on the data shown in Figure 1, Reference 7 should be consulted.

THE MINIMUM PILOT RATING ANALYSIS METHOD

The rating expression discussed previously is, in fact, based on a single data source; namely, the results in Reference 3. Direct correlation with other V/STOL hover data is impossible because of a general lack of measured pilot parameters and closed-loop performance under gust disturbances.

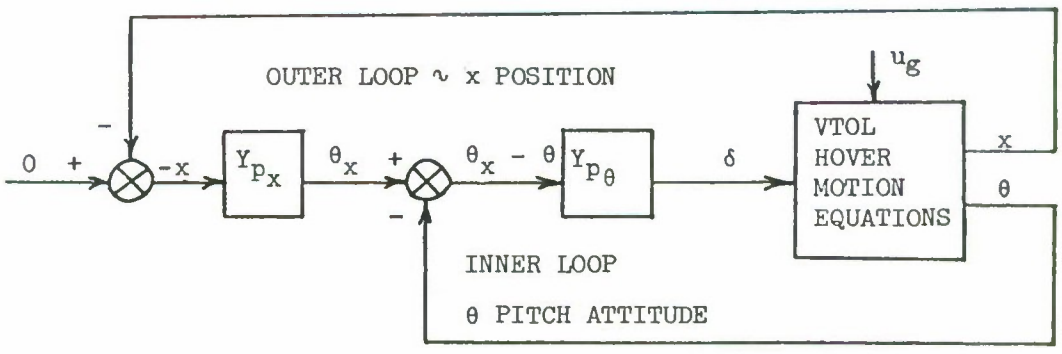
On the other hand, if pilot parameters and closed-loop performance <u>could be predicted</u>, a number of other data sources might be used to further validate

the rating expression. Although general hover trends have been predicted before (Reference 2), the current method requires quantitative predictions if quantitative comparisons with pilot rating are to be made. The following material describes how such a prediction analysis method was evolved

Optimal Pilot Concept: The idea that the human operator adjusts his method of control to produce low frequency performance which ".... is optimum in some sense analogous to that of minimum mean-squared tracking error" (Reference 8) is not new. This concept is also employed in the closed-loop analysis procedure that has been used in many applications (e.g., Reference 2) along with other factors such as gain sensitivity, phase margin, etc. However, in only a few cases has a fixed-form pilot model been adjusted solely on a minimum error basis (e.g., References 9 and 10). This approach is attractive, however, since it is in no way "arty".

Free-form (i.e., no assumed pilot form) optimal pilot models have also been successful (Reference 11 and the references therein). Nonetheless, in each case the question of which performance "error" should be minimized presents the biggest unknown. Fortunately, the rating expression developed above suggests a natural performance index to minimize. That is, since pilot ratings appear to be directly related to σ, it seems natural to assume, as an extension of simpler task results, that $\sigma = \sigma_x + 10\sigma_q$ is to be minimized.

To investigate this concept, a simplified version of the fixed-form pilot model of Reference 3 was implemented on an analog computer along with the hover dynamics and gust simulation. Since this pilot model form was used directly to compute the pilot parameters used in Figure 1, its use again was natural. The loop closures with pilot models were in the manner of Reference 3 as represented in the sketch below.



where: $Y_{p_x} = K_{p_x}(T_{L_x} s + 1)$

$Y_{p_\theta} = K_{p_\theta}(T_{L_\theta} s + 1)e^{-\tau s}$; $\tau = 0.44$ seconds

Figure 2 shows the variation in σ as a function of each of the four variable pilot model parameters, with the remaining three parameters held constant at "nominal" values. The case in question is PH 5, of Reference 3. Also shown in the figure are: a) optimal, in a minimum σ sense, parameters ("opt"), and b) actual measured pilot parameters (pilot). The pilot gain K_{p_θ} is not shown because it represented an unstable situation with the other three parameters at their nominal values.

The next question is the selection of a set of four pilot parameters that form a "prediction" of how the pilot operates. Note that:

a) Measured pilot gain, $-K_{p_x}$, seems to fall short of the "opt" value by a small amount, perhaps representing a finite stability gain and phase margin (per adjustment rules in Reference 8). This reduced value of gain has little effect on performance.

b) Measured pilot pitch lead, T_{L_θ}, is also less than the optimal value. In this case, a fairly large T_{L_θ} performance penalty is evident due to the "off optimal" condition.

c) Measured pilot displacement lead, T_{L_x}, is slightly larger than optimal, but the performance variation with lead T_{L_x} is very small in this region.

Based on these observations, the initial prediction ground "rules" emerged as follows:

a) Select pilot gains that produce minimum values of σ with a reasonable gain margin (i.e., a small increase in gain will not produce excessively large σ values). In general, keep the gain as low as possible while still producing "good" performance.

b) Select pilot lead terms as the optimal, if greater than 0.2 seconds, or as the smallest value that produces acceptable performance. The former evolved from an earlier rating expression that had a "dead-zone" in rating increment up to 0.2 seconds. This was later removed for simplicity.

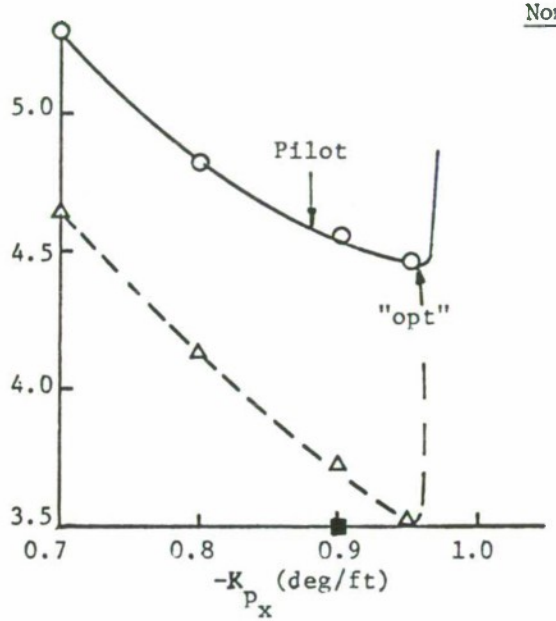
Using the somewhat hazy rules, the "predicted" set of pilot parameters shown in Figure 2 was used to predict performance, rating, and approximate closed-loop bandwidths (ω_{c_x} and ω_{c_θ}). The latter were obtained from a strip chart recording and served mainly to validate the analog simulation. A digital computer was also used to validate the analog computer performance (σ) values. The resulting values are shown on Figure 2, along with actual measured values.

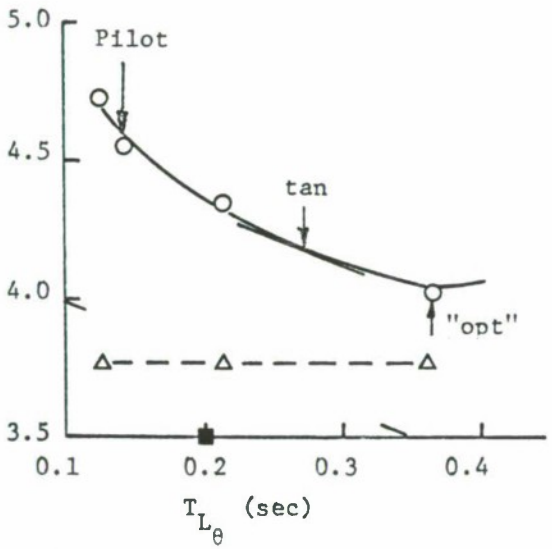
Even though the "selection rules" are somewhat arbitrary, and the resulting selected pilot parameters do not exactly match measured values, the overall rating, performance, and pilot parameter values compare very well with measured data.

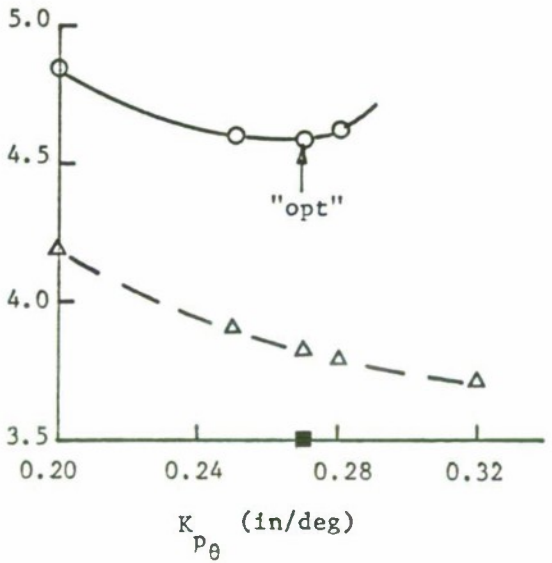
Minimum Rating Concept: Although the results shown in Figure 2 compare well with measured data, the selection of pilot lead terms remains "arty". Returning to the rating expression developed earlier, one could interpret the

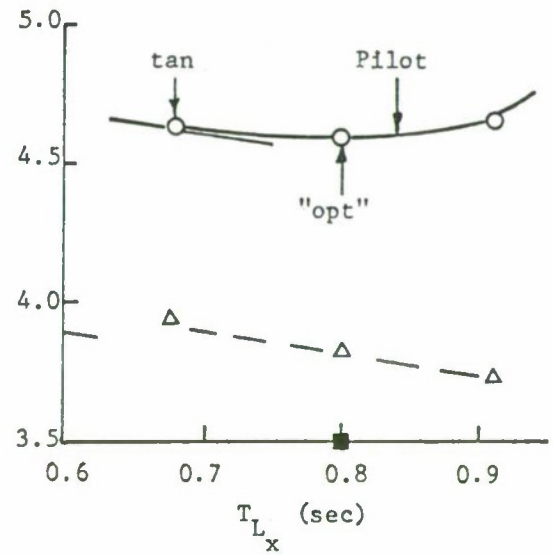
Nominal Values: $K_{p_\theta} = 0.28$ $-K_{p_x} = 0.88$ O—O $(\sigma_x + 10\sigma_q) = \sigma$

$T_{L_\theta} = 0.14$ $T_{L_x} = 0.84$ Δ—Δ σ_x









σ_x = Feet σ_q = RAD/Sec

■ = Predicted Value

σ_{u_g} = 5.1 FT/SEC

	PREDICTED	MEASURED
T_{L_θ}	0.20	0.14
$-K_{p_x}$	0.90	0.88
$K_{p\theta}$	0.27	0.38
T_{L_x}	0.78	0.84
σ_x	3.89	3.59
σ_q	0.055	0.064
ω_{c_θ}	2.33	2.40
ω_{c_x}	0.89	0.76
PILOT RATING	4.78	5.0

Figure 2. Initial Prediction Attempt

expression as a mathematical statement of the pilot's desire to "achieve adequate performance (small σ) with a minimum of effort (small T_L's)". If this were indeed true, then a trade in T_L's vs. performance , σ, should be evident. That is, an increase in T_{L_θ} or T_{L_x} is "justified" only if the "pay off" in terms of better performance (σ) is great enough. Stated another way, the pilot may attempt to minimize his own rating of the vehicle dynamics by adjusting his parameters to minimize R.

This interesting, but unproven, concept was informally presented to several pilots. Most accepted the idea as quite reasonable. In any event, the concept can be used to provide a rigid method of selecting pilot leads. If no change in rating results (i.e., an increase in T_{L_θ} increases R_2, but also improves performance thereby reducing R_1 by a similar amount) then ΔR is zero, and this defines a critical slope to establish an "optimal" (minimum rating) value of T_{L_θ}.

These "critical" slopes are shown in Figure 2. The tangent points (marked tan.) with the true σ vs. T_L curves determine the predicted value. It can be seen that at least reasonable pilot leads are selected using this procedure, and it was decided to use the minimum rating and critical slope concept for all subsequent predictions.

In summary, the pilot-vehicle analysis method suggested above uses a fixed-form pilot model with pilot leads selected to give a minimum (best) pilot rating of the vehicle dynamics. Pilot gains are also selected to provide good closed-loop performance (and minimum rating) with a "reasonable" gain margin. In conjunction with the rating expression it has the unique capability, for the hover task, to allow the prediction of pilot parameters, closed-loop pilot-vehicle performance under gusty conditions, and pilot opinion rating. The only "arty" aspect is pilot gain selection, and that will also be "automated" in a subsequent section.

CORRELATION OF PREDICTED RATINGS WITH DATA FROM VARIOUS SOURCES

Before temporarily leaving the data bank in Reference 3 (upon which the rating expression was developed) another correlation is of interest. That is Case PH 3 which was also evaluated with the addition of a first order actuator simulation using different actuator time constants. Reference 3 does not include measured pilot parameters for the non-zero time constant cases, but the minimum rating prediction method may be used to evaluate the effects of the non-ideal actuator. The results were as follows: (τ_c = actuator time constant);

Case	τ_c(sec)	Pilot Rating	Completely Predicted Pilot Rating
PH 3	0.10	4.0	4.14
PH 3	0.50	6.0	5.93

This rather amazing correlation speaks for itself. The value of this approach in evaluating the effects of control system dynamics is also obvious.

The minimum rating method was next used to predict pilot rating for various hover dynamics as presented in some seven separate references that cover fixed-base simulations, moving-base simulations, variable stability helicopter tests, and actual flight tests. Detailed results are presented in Reference 7. However, most correlations with actual pilot ratings have been very good to excellent.

THE "PAPER PILOT"

The predicted rating results discussed so far were obtained using the analog computer. The minimum pilot rating was found by manually searching for the minimizing pilot leads and selecting pilot gains to provide a "reasonable" gain margin. It became clear that a fully automatic scheme was required to avoid the manual drudgery of exercising the model on the analog computer. A fully automatic scheme would also remove the last remnants of "artiness" from the procedure. This led to a digital computer program, dubbed the "Paper Pilot", for predicting pilot rating in the hover mode.

A precise mathematical model for predicting pilot rating of the flying qualities of a VTOL aircraft in the hover mode was developed. The model includes the following important elements:

1. The longitudinal equations of motion for the VTOL aircraft flying in the hover mode.

2. A stochastic input model representing gust disturbances to the aircraft.

3. A fixed form pilot model with four free parameters. The parameters include two pilot gains and two pilot leads.

4. A cost functional similar to the expression for R which is made up of a measure of hover performance and pilot workload. For a given aircraft configuration and gust intensity, the cost functional is a real valued function of the pilot parameters.

The scheme is to first find the pilot parameters which minimize the cost functional. The minimizing gains are tested for a 20% stability margin. If this test fails, the pilot gains are adjusted to provide a 20% stability margin. Otherwise the minimizing gains are used. The resulting pilot parameters and closed-loop performance are then used in the expression for R to compute a "Paper Pilot" rating of the handling qualities of the VTOL aircraft in the hover mode.

For a detailed description of the mathematical model and the digital computer program, see Reference 12.

A typical computer printout is shown in Figure 3. Vehicle dynamics, gust intensity, convergence criteria, and an initial guess of the pilot parameters represent the input data. The value of the cost functional, pilot parameters, and closed-loop performance values are all printed out for each iteration in the minimizing routine. Pilot gains are adjusted for 20% stability margin, if

VTOL Hover Task AFFDL-TR-67-152

CASE -- PH36

```
MU          =    0.02484
XU          =   -0.36000
MQ          =   -1.00000
MTHETA      =    0.
MDELTA      =    0.27900
SIG GUST    =    2.10000

ESTIMATE    =    2.00000
EPSILON     =    0.20000
LIMIT       =   15

KPTH        =    0.35199
TLTH        =    0.40158
KPX         =   -1.89088
TLX         =    0.01874
```

VEHICLE DYNAMICS

GUST INTENSITY

ITERATION DATA

PILOT PARAMETER "GUESS"

CLOSED LOOP PERFORMANCE

PILOT RATING	PILOT PARAMETERS				σ_q	σ_θ	σ_u	σ_x	
PR	KPTH	TLTH	KPX	TLX	SIGQ	SIGTH	SIGU	SIGX	//G//
3.46934	0.35199	0.40158	-1.89088	0.01874	3.84784	2.73143	1.06992	1.28626	1.42919
3.45810	0.36634	0.41396	-1.89290	0.01892	3.79603	2.64672	1.03319	1.26139	1.20794
3.45385	0.36693	0.43052	-1.89680	0.01429	3.58671	2.60570	1.03409	1.26508	0.19198
3.45283	0.36998	0.43050	-1.90140	0.01794	3.63575	2.59350	1.02119	1.25285	0.27737
3.44699	0.35835	0.42899	-1.93360	0.03192	3.66613	2.61960	1.01770	1.23470	0.36780
3.44512	0.36345	0.42903	-1.95098	0.03917	3.77175	2.60558	0.99491	1.20869	0.74713
3.44292	0.36033	0.43175	-1.95211	0.03613	3.69717	2.61072	1.00539	1.21714	0.11925

ITERATIONS

EXIT CONJUGATE GRADIENT SEARCH WITH IER = 0

3.44081	0.36674	0.43826	-1.95924	0.02947	3.66164	2.59582	1.00417	1.21397	0.40506

PILOT GAINS ARE ADJUSTED TO PROVIDE 20 PER CENT STABILITY MARGIN. PILOT RATING IS MODIFIED AS FOLLOWS

3.47260	0.34955	0.43826	-1.86740	0.02947	3.41610	2.57328	1.03056	1.28222	← FINAL RESULTS

GUST REGION CODE IS 111

(3.5)	(0.91)	(0.42)	(-1.28)	(0.08)	(2.26)	(1.95)	(0.9)	(1.51)	← PILOT B MEASURED VALUES

FOR A GUST DISTURBANCE MODEL WITH A STANDARD DEVIATION OF 2.100, THE PAPER PILOT RATES THIS AIRCRAFT

FINAL NUMERICAL RATING

3.47260

IN THE PRECISION HOVER MODE.

FLYING QUALITIES ARE CLEARLY ADEQUATE FOR THE MISSION FLIGHT PHASE OF PRECISION HOVER (PH).

"VERBAL" EVALUATION

******* LOOK -- UP IN THE SKY. IT IS A BIRD. IT IS A PLANE. NO -- IT IS PAPER PILOT. *******

Figure 3. Typical Computer Printout

necessary, and the "paper pilot" rating is computed and printed out. A verbal evaluation of the flying qualities is given based on the numerical rating.

A typical computer run with 15 iterations takes around 80 seconds.

"PAPER PILOT" COMPUTATIONAL RESULTS

The "paper pilot" program was used to compute the "paper pilot" rating for 26 of the aircraft configuration-gust intensity combinations taken from Reference 3. A detailed summary of the results is presented in Reference 12. A summary of these results is given in Table 1. The actual pilot data is from Reference 3. Data was obtained from fixed-base simulation using two pilots,

Data	Trend	Mean of the Difference*	Standard Deviation of the Difference*
Pilot Rating	For M_q=-3 or -5 "paper pilot" is too low For M_q=-1 "paper pilot"PR is too high except for cases PH35 and PH36 "Paper Pilot" tends to rate too low with increasing σ_{u_g}	.12	.55
K_{p_θ}	"Paper Pilot" K_{p_θ} is lower except in 2 cases	.17	.14
T_{L_θ}	"Paper pilot" T_{L_θ} is higher except in 3 cases	-.09	.10
K_{p_x}	"Paper pilot" K_{p_x} is higher in every case	-.90	.46
T_{L_x}	"Paper pilot" T_{L_x} is lower in every case	.26	.21
σ_θ	"Paper pilot" σ_θ is greater in every case	-.61	.31
σ_q	"Paper pilot" σ_q is greater except in one case	+1.09	.62
σ_x	"Paper pilot" σ_x is less except in 5 cases	.32	.45
σ_u	"Paper pilot" σ_u agrees well with actual σ_u	-.12	.18

*The difference is taken to be (actual pilot data)-("paper pilot" data). In those cases where two actual pilots were used, the average is used in the computation of the differences. Does not include the two cases where $\sigma_{u_g} > 12$.

TABLE 1

COMPARISON OF "PAPER PILOT" RESULTS TO ACTUAL PILOT RESULTS

Two trends are apparent in the pilot rating data. The first trend is with respect to M_q. For M_q = -3 and -5, the "paper pilot" rating tends to be too low and for M_q = -1 the "paper pilot" ratings tend to be too high. The second trend in the pilot rating results is in respect to σ_{u_g}. The "paper pilot" rating tends to be too low with increasing gust intensities. For the two cases where the gust intensities are 20.6, the "paper pilot" rating is a full rating point below the actual pilot rating. For this reason, the model, in its present form, is not considered valid for $\sigma_{u_g} > 12$.

These two trends suggest that a better fit to the actual pilot ratings might be achieved by an increase in the weighting on performance and a corresponding decrease in the weighting on T_{L_θ}. Such a modification may also result in a model which is valid for higher gust intensities.

Furthermore, since the "paper pilot" had poorer performance in pitch rate than the actual pilots, an increase in weighting on σ_θ and σ_q might also result in better agreement between paper pilot performance. This alteration to the model is consistent with the changes suggested for improving the agreement between "paper pilot" ratings and the actual pilot ratings.

The agreement between the "paper pilot" T_{L_θ} and the real pilot matched T_{L_θ} is considered good. However, T_{L_x} does not compare as well. The former is, however, more important in respect to pilot rating.

In general the "paper pilot" ratings agree well with the real pilot ratings. A standard deviation of .55 in the differences is considered comparable to the variations in pilot ratings from one pilot to another and even to the variations in rating for the same pilot in repeated runs. In addition, the results indicate that the "paper pilot" can be "fine tuned" to give even better agreement with actual pilot ratings, as well as fair agreement with actual performance and matched pilot parameters.

The "paper pilot" was used in conjunction with describing function theory to analyze a nonlinear case (Reference 13). The nonlinearity considered is limited rate authority in the Stability Augmentation System (SAS). The pilot rating was computed for three authority limits and the results are as follows:

Rate SAS Authority M_{RA}, rad/sec^2	Pilot Rating	"Paper pilot" Rating
Unlimited	3.0	2.9
.1	4.0	4.6
.03	5.5	5.7
0	6.0	5.8

The actual pilot ratings are obtained from Reference 3. Clearly, the agreement is excellent.

APPLICATION TO FLYING QUALITIES SPECIFICATION

As noted in the introduction, the motivation for this work was specification oriented, and Appendix B of Reference 7 presents a proposed hover specification using the development. In short, it allows the contractor to predict pilot rating using the rating expression and the minimum rating approach. Other limits on pilot leads and gain margins are included. The resulting predicted ratings must then meet certain minimum requirements. This approach has a number of advantages in that it:

a. does not require an assumed type of augmentation (i.e., rate, attitude, etc., systems);

b. accounts for control system and other higher order dynamics;

c. accounts for the effects of gust intensity (at least up to about 12 ft/sec rms); and

d. is extremely simple to use.

SUMMARY AND CONCLUSIONS

An examination of the correlation of pilot rating with closed-loop performance and pilot model parameters has led to: a) a method of calculating pilot rating from measured data for the hover task; b) a new approach to pilot-vehicle analysis with potentially wide application that can provide "automatic" prediction of pilot model parameters, closed-loop performance, and pilot rating; and c) an entirely new approach to the evaluation and specification of flying qualities. These developments have been tested for only one task, VTOL hover. However, there is every indication that much wider application is possible both in terms of other tasks and uses.

The results reported here are far too good to represent an "accident". Perhaps the gap between flying qualities and human response theory is now a lot smaller. Or, maybe the hover task is a rare exception, and what now appears to be a widely applicable approach is, instead, of limited value. In any event, the authors feel that the art of specifying and evaluating flying qualities will never be quite the same again.

REFERENCES

1. Teper, G. L., and Jex, H. R., "Synthesis of Manned Booster Control Systems Using Mathematical Pilot Models", Systems Technology, Incorporated, Sixth Annual Symp. Prof. Group on HF in Electronics, IEEE, 6-8 May 1965.

2. Craig, S. J., and Campbell, A., "Analysis of VTOL Handling Qualities Requirements", Part I, Systems Technology, Inc., AFFDL-TR-67-179, October 1968.

3. Miller, D. P. and Vinje, E. W., "Fixed-Base Flight Simulator Studies of VTOL Aircraft Handling Qualities in Hovering and Low-Speed Flight", United Aircraft Research Laboratories, AFFDL-TR-67-152, January 1968.

4. Vinje, E. W., and Miller, D. P., "Analytical and Flight Simulator Studies to Develop Design Criteria for VTOL Aircraft Control Systems", United Aircraft Research Laboratories, AFFDL-TR-68-165, April 1969.

5. McDonnell, John D., "Pilot Rating Techniques for the Estimation and Evaluation of Handling Qualities", Systems Technology, Inc., AFFDL-TR-68-76, December 1968.

6. Adams, J. J., "An Approach to the Determination of Aircraft Handling Qualities Using Pilot Transfer Functions", presented at the Fourth NASA Inter-Center Control Systems Conference, November 4-5, 1969, Cambridge, Mass.

7. Anderson, R. O., "A New Approach to the Specification and Evaluation of Flying Qualities", Air Force Flight Dynamics Laboratory, AFFDL-TR-69-120, November 1969.

8. McRuer, D. T., et al, "Human Pilot Dynamics in Compensatory Systems", Systems Technology, Inc., AFFDL-TR-65-15, July 1965.

9. Wanamaker, J. E., and Sower, W. A., "Extension of Pilot Describing Functions to Multiple Compensatory Tracking Tasks", GE/EE/69-18, Air Force Institute of Technology, March 1969.

10. Mannen, J. T., and Duggar, L. C., "Development of Human Describing Function Models for Nonlinear Control Elements", GE/EE/69-6, Air Force Institute of Technology, March 1969.

11. Elkind, J. I., et al, "An Optimal Control Method for Predicting Control Characteristics and Display Requirements of Manned-Vehicle Systems", Bolt Beranek and Newman, Inc., AFFDL-TR-67-187, June 1968.

12. Dillow, James D., "The 'Paper Pilot'--A Digital Computer Program to Predict Pilot Rating for the Hover Task", Air Force Flight Dynamics Laboratory, AFFDL-TR-70-40, January 1970.

13. Personal communications with "Paper Pilot", January 1970.

THE APPLICATION OF HUMAN OPERATOR DESCRIBING FUNCTION THEORY TO THE PREDICTION OF TRACKING PERFORMANCE IN THE CHEYENNE SWIVELING GUNNER'S STATION

P. Briggs
Avionic Controls Department
General Electric Company
Binghamton, New York

L. G. Hofmann
Systems Technology, Inc.
Princeton, New Jersey

12 January 1970

ABSTRACT

The Cheyenne aircraft is a sophisticated fire support system having such diverse missions as the escort of troop carrying helicopters and the delivery of guided missiles against heavy point targets.

The heart of the Cheyenne fire control system is the General Electric Swiveling Gunner's Station, by means of which the gunner tracks moving ground targets in the presence of ownship attitude and position variations. Since the control of ballistic weapons and the guidance of missiles in relation to the line-of-sight is automatically accomplished, tracking accuracy is the fundamental performance parameter of the Swiveling Gunner's Station.

Because of the stringent requirements of the Cheyenne weapon system, a tracking accuracy prediction during the development phase was mandatory. The application of human operator describing function theory to this prediction problem is the subject of the present paper.

The most significant feature of the paper is a comparison of a gunner model obtained from experimental tracking data with an alternate form of the extended crossover model. Measurements of gunner remnant are also compared with a recently published controller remnant model.

THE APPLICATION OF HUMAN OPERATOR DESCRIBING FUNCTION THEORY TO THE PREDICTION OF TRACKING PERFORMANCE IN THE CHEYENNE SWIVELING GUNNER'S STATION

INTRODUCTION

BACKGROUND - THE CHEYENNE WEAPON SYSTEM

The Cheyenne aircraft is a rigid-rotor, compound helicopter with a top speed of over 200 knots. Its armament includes machine guns, rockets, and TOW anti-tank missiles. The two-man crew consists of the pilot and the copilot/gunner. A view of the aircraft in flight is given in Figure 1.

Since May, 1966, the Avionic Controls Department of the General Electric Company (GE/ACD) located at Binghamton, New York, has been under contract with the Lockheed-California Company for the design and fabrication of the Cheyenne Swiveling Gunner's Station.

THE SWIVELING GUNNER'S STATION

The heart of the Cheyenne weapon control system is the General Electric Swiveling Gunner's Station (SGS). With it the copilot/gunner is able to acquire and track targets and select and fire his weapons. A "gunner's eye view" of the SGS is given in Figure 2.

The most essential performance parameter of the SGS is the accuracy with which the gunner can track a moving target in the presence of aircraft attitude variations and translations. The gunner is aided in his tracking task by the periscope optical system (3x and 12x magnification levels available), by the 2-axis stabilized sighthead of the SGS, and by computed linear motion compensation (LMC) commands which are transmitted from the aircraft computer central complex. Ideally, the combination of space stabilization and LMC will cause the line-of-sight (LOS) to remain on a fixed ground target in the face of aircraft rotations and translations. Accordingly, the tracking task of the gunner consists of nullifying any small, residual errors present in the LMC and commanding the needed rates to maintain the LOS on a moving ground target. Manual control of the LOS in elevation and train is by means of a spring-restrained, thumb-force button located at the top of the fixed right hand-grip (see Figure 2).

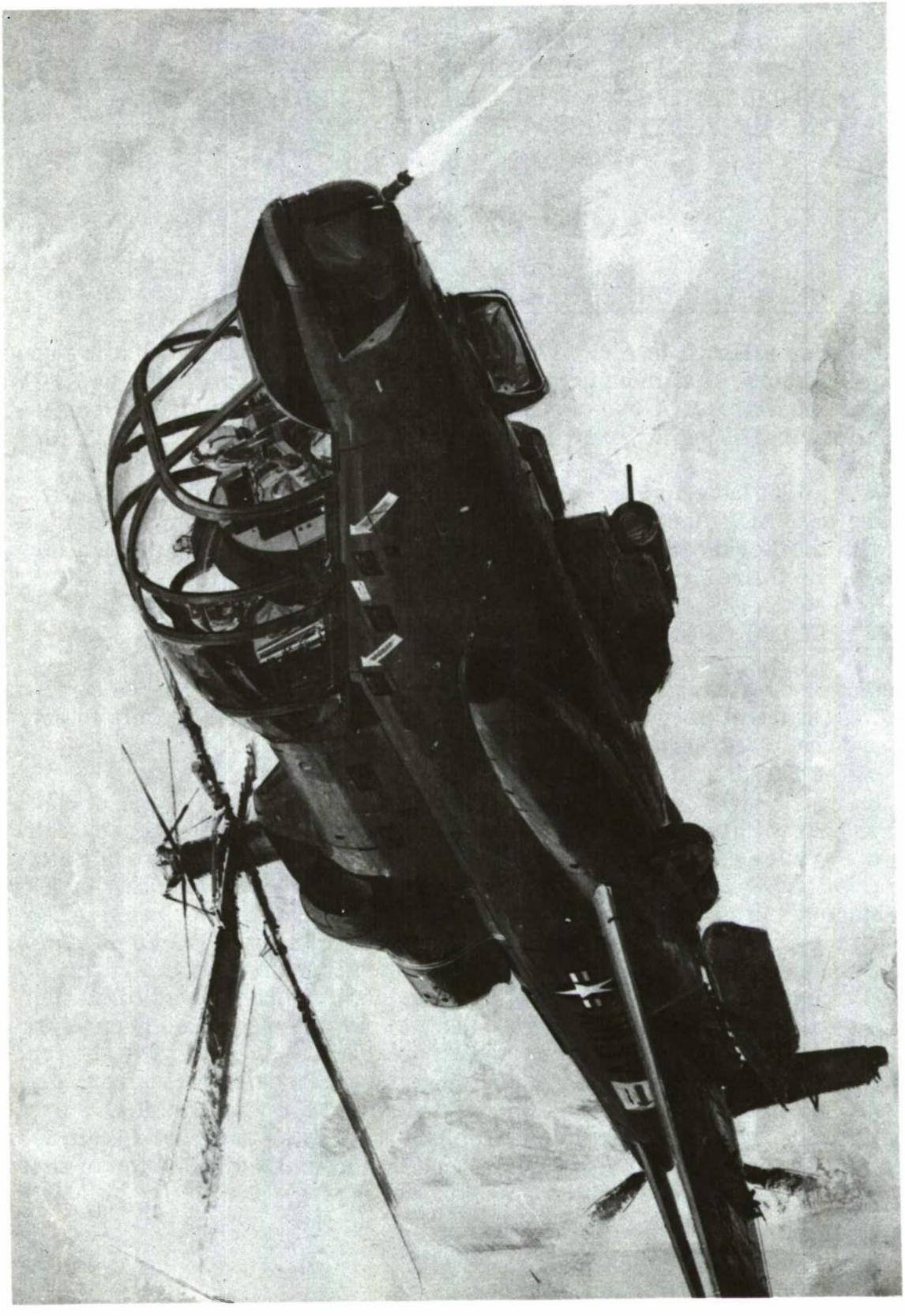


Figure 1. The AH-56A Cheyenne Aircraft in Flight

The actual position of the LOS in space is determined by the orientation of the mirror, which is part of the stabilized sighthead and may be seen in Figure 3. One may visualize the gunner's seat and the mirror as moving essentially together in response to train commands. In actual fact, the sighthead has its own train servo which has a bandwidth substantially higher than that of the seat drive; and it can move independently of the seat for small angles of relative motion. Thus, in tracking, the sighthead servos actually generate the entire inertial LOS rate in both elevation and train, and the seat drive follows so as to null the angle between it and the LOS about the train axis.

THE FREQUENCY DOMAIN ERROR ANALYSIS

At the very outset, the need for a careful analysis of tracking accuracy was recognized. Accordingly, GE/ACD began the formulation of two computer simulations of the Cheyenne weapon control system. Both of these were directed primarily toward the prediction of the LOS error for a variety of engagements, and therefore they both included a model of the gunner's tracking behavior. However, they differed with respect to the domain in which the error generation and propagation characteristics of the weapon control system were modeled. In the first, the system error generation and propagation model was formulated in the time domain and its outputs included the LOS error time histories resulting from the introduction of individual sensor error values. In the second, the error generation and propagation characteristics of the system were modeled in the frequency domain, its final outputs being the spectral densities of the elevation and train components of the LOS error.

This paper is concerned with the application of human operator describing function theory to the prediction of the LOS error in the frequency domain. In fact, the accurate modeling of the human gunner is regarded as the most crucial aspect of this prediction problem.

A block diagram showing the organization of the frequency domain error analysis is given in Figure 4. The primary functions of its four main sections may be summarized as follows:

The Time Domain Encounter Model - In this section the kinematics of an engagement of the Cheyenne aircraft and a moving ground target are simulated in considerable detail. Figure 5 is an illustration of such an engagement. In particular, the simulation includes purposeful aircraft maneuvers, random aircraft attitude variations due to wind gusts, and a target motion model consisting of random longitudinal, lateral, and vertical accelerations superimposed on a nominal, steady velocity vector. The essential outputs of this section of the error analysis are the time histories of all the system parameters which influence the generation and propagation of errors in the Cheyenne weapon control system.

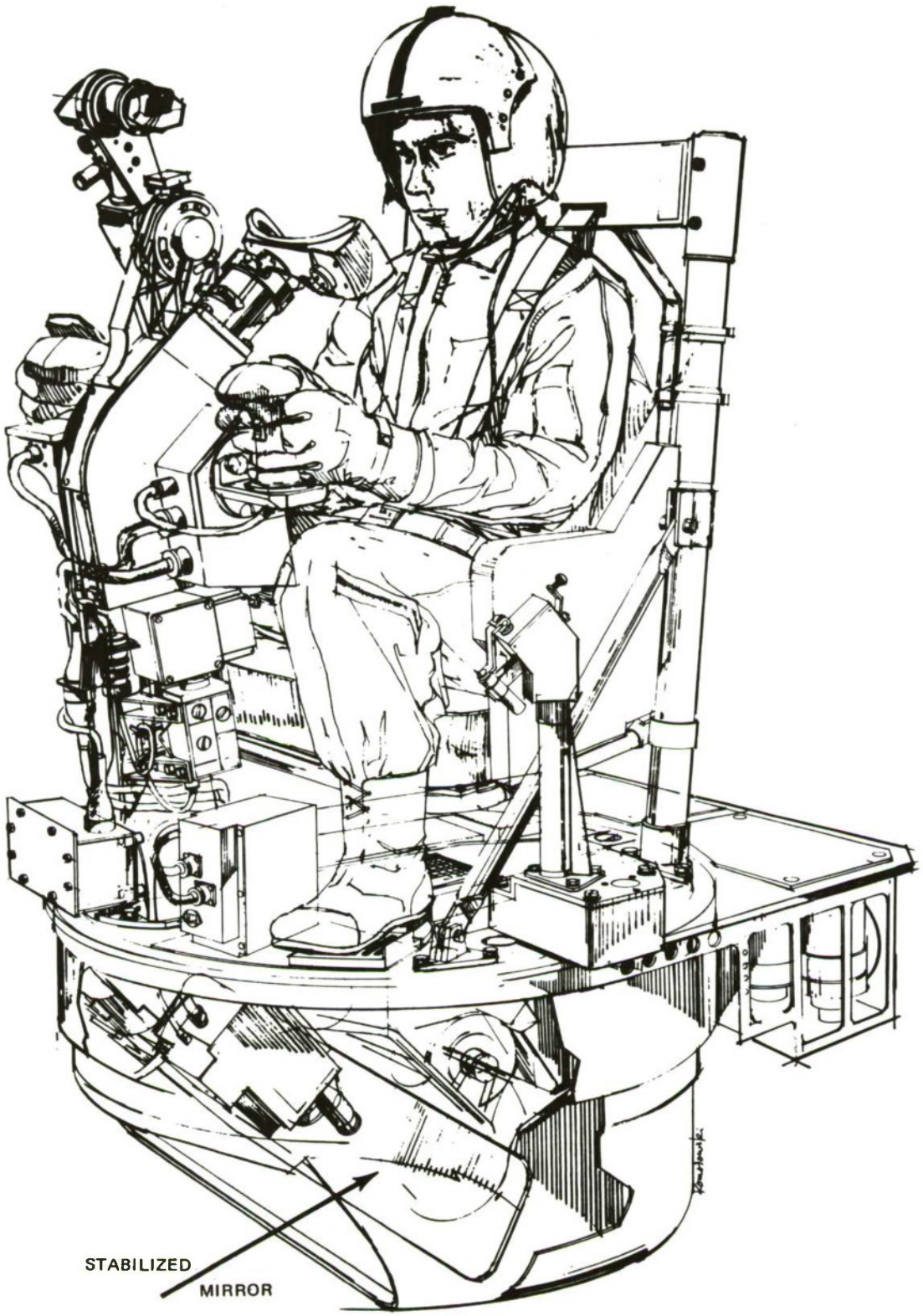


Figure 3. An Overall View of the Swiveling Gunner's Station

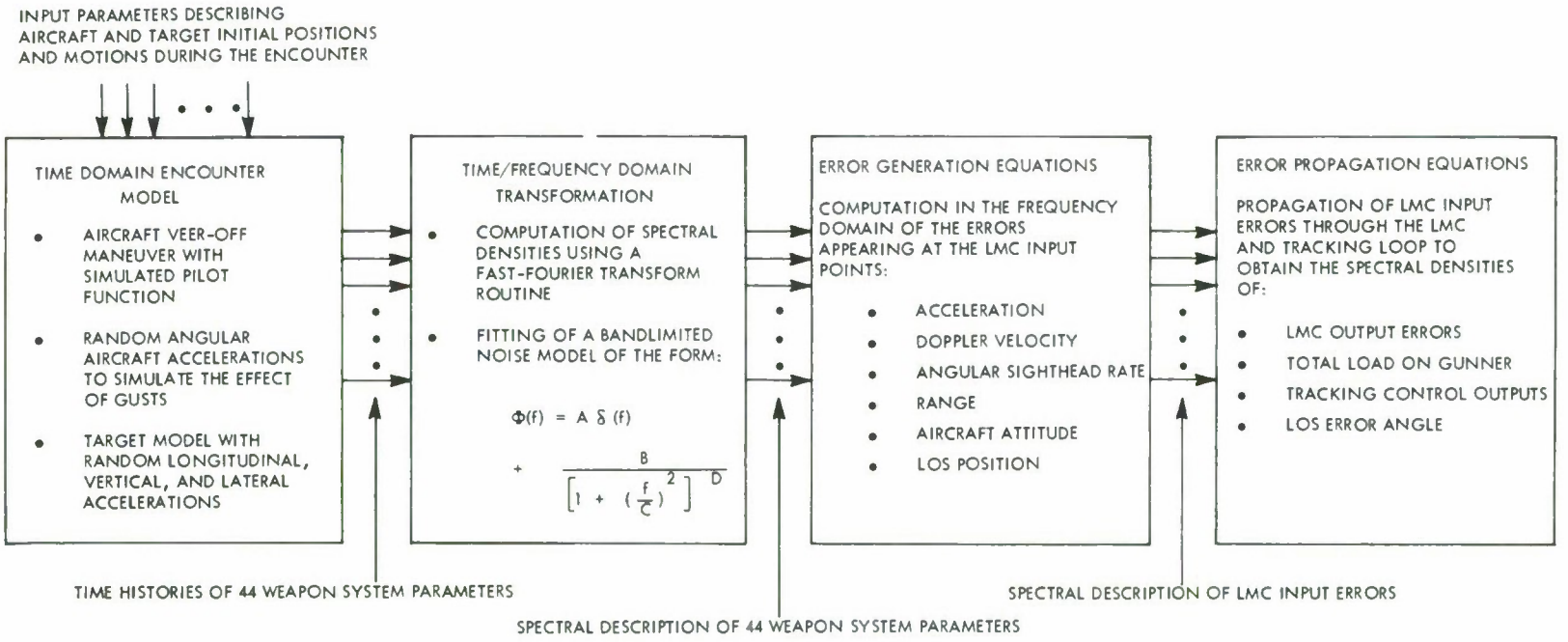


Figure 4. A Block Diagram of the Frequency Domain Error Analysis

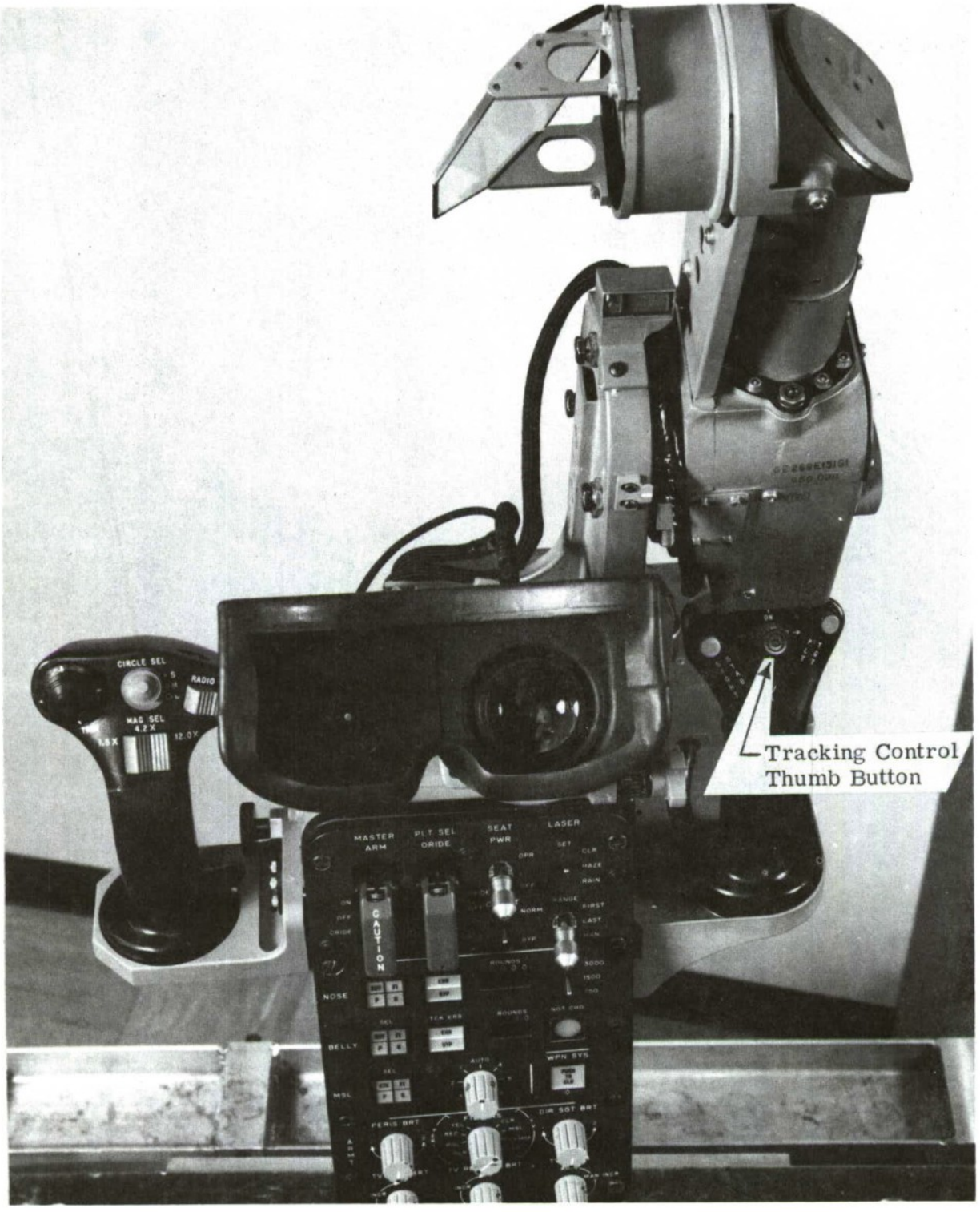


Figure 2. A "Gunner's Eye View" of the Cheyenne Swiveling Gunner's Station

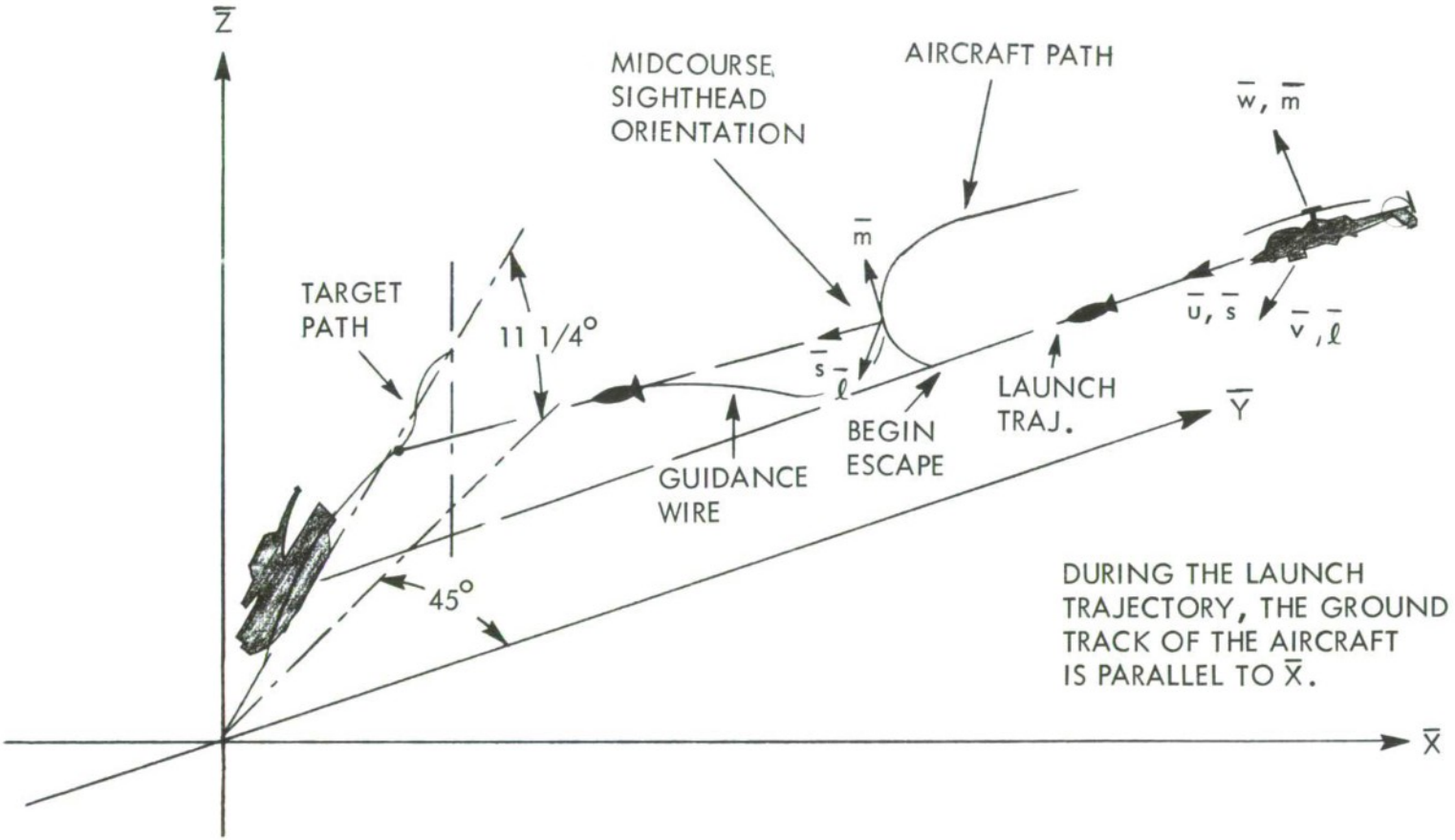


Figure 5. An Illustration of an Engagement of the Cheyenne Aircraft with a Moving Ground Target

The Time to Frequency Domain Transformation - In this section the time histories generated in the encounter model are transformed into spectral densities.

The Error Generation Equations - The generation of errors at the input points to the LMC is modeled in this section. The spectral densities of the system parameters (or their auto-correlation functions, as the need may demand) are combined and transformed to yield the spectral densities of these errors.

The Error Propagation Equations - In this fourth and last section the propagation of errors through the LMC computations and the tracking loop is modeled. The final outputs of the error analysis are the spectral densities of the elevation and train components of the following:

- The LMC error
- The total effective angular rate disturbance to the tracking loop
- The gunner's hand control ouputs
- The angular error in the LOS.

The error propagation characteristics of the tracking loop in which the gunner functions are the main concern of this paper. Given a particular encounter situation and LOS stabilization aids, it is the effectiveness of the gunner in this closed-loop tracking task which is the main determinant of the system tracking accuracy. In Figure 6, a block diagram of the closed-loop tracking task is shown.

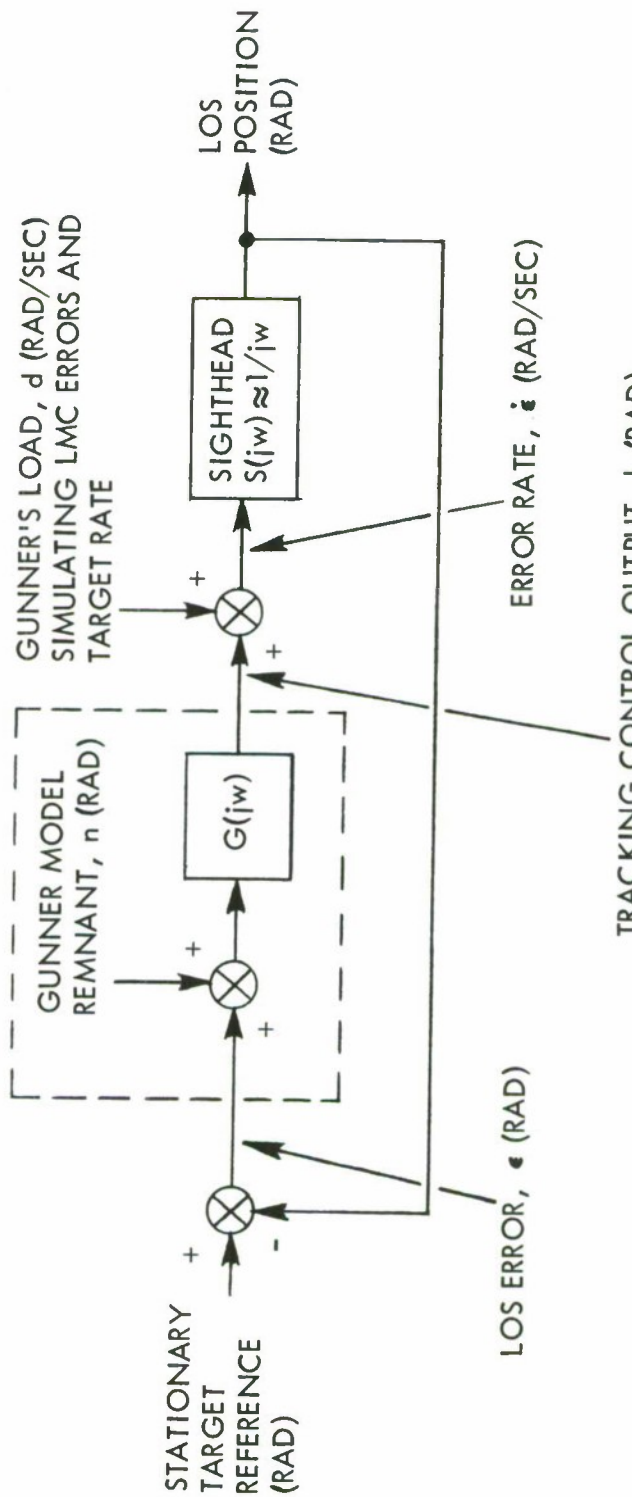


Figure 6. A Block Diagram of the SGS Tracking Loop

Section I

DESCRIPTION OF THE SGS TRACKING LOOP

THE SIGHTHEAD RESPONSE FUNCTION

As shown in Figure 6, the sighthead response, $S(j\omega)$, is approximated by $(j\omega)^{-1}$. To be precise, its departure from this ideal response may be represented by a simple lag at 85.7 radians/second or 13.6 Hz. This may be compared with the measurement bandwidth of the describing function data to be presented in the sequel, which was 7.58 Hz. It also may be compared with the tracking loop crossover frequency, which is in the neighborhood of .75 Hz.

Because the bandwidth of the sighthead is sufficiently high so as to not materially affect the LOS error resulting from a given disturbance input, it will be neglected in the sequel. That is, the assumption will be made that the sighthead response is adequately represented by $(j\omega)^{-1}$ throughout the frequency range relevant to tracking error prediction.

THE GUNNER'S TRACKING BEHAVIOR

Two key assumptions are made regarding the gunner's behavior in the tracking loop, which in turn are reflected in the structure of the tracking loop diagram shown in Figure 6. These assemptions are:

- He is assumed to respond in a compensatory fashion to the LOS error, ϵ.
- His elevation and train responses are assumed to be uncoupled.

Accordingly, Figure 6 is regarded as pertaining to either the elevation or train axis, the gunner's describing function being denoted $G(j\omega)$ and the response of the pertinent sighthead servo being denoted $S(j\omega) = (jw)^{-1}$.

The first assumption deserves further comment. It must be appreciated that the field-of-view afforded the gunner, even with 12x periscope magnification, is such that his display is pursuit in nature. That is, the reference input to the tracking loop, the LOS position, and the LOS error are all, in effect, separately displayed to the gunner. However, the gunner's behavior is as if the display were compensatory; that is, as if only the LOS error were displayed. This is a consequence of the sighthead dynamics. To the

gunner, the sighthead appears to have essentially pure rate dynamics, and it has been shown in carefully controlled laboratory experiments (Ref. 1) that when controlling pure rate dynamics a human subject adopts compensatory behavior, regardless of whether the display is compensatory or pursuit. The reason for this is twofold: 1) In the case of pure rate dynamics, pursuit behavior affords no performance improvement over compensatory; and 2) Pursuit behavior is more demanding of the subject.

THE GUNNER'S LOAD AND THE ERROR RESPONSE RATIO

This is an appropriate point to introduce definitions of three terms which will be mentioned frequently in the sequel. First, the gunner's load is defined as the total effective angular rate disturbance to the tracking loop, designated d in Figure 6. This load combines the angular LOS rates demanded by motion of the target and the disturbances to the LOS arising from LMC errors. Accordingly, the reference input to the tracking loop is regarded as a stationary target position since the effect of target motion is included in the load, d.

Next, the overall error response ratio, η, is defined as the standard deviation of the LOS error, σ_ϵ, divided by the standard deviation of the load, σ_d. That is,

$$\text{Overall Error Response Ratio} = \eta = \sigma_\epsilon / \sigma_d \tag{1}$$

Whereas the overall error response ratio, η, reflects the performance of the manual control system (consisting in this case of the gunner, his tracking control, and the sighthead servo), the load, d, reflects the level of the inputs forcing the tracking loop in which the gunner is functioning. In fact, η, is indicative of tracking system effectiveness in attenuating the disturbing influence of the load.

In the absence of remnant, the basic concept of the overall error response ratio can be readily extended to express the error response ratio obtained over a certain frequency interval. By this means, the effectiveness of the tracking system in attenuating error in each of several adjacent frequency intervals can be evaluated. In the limit, as the intervals are made to approach zero, the error response ratio, $E(\omega)$, which is a continuous function of ω, results.

This function is simply related to the gunner describing function, $G(j\omega)$, and the sighthead response function, $S(j\omega) = (j\omega)^{-1}$.

$$E(\omega) = \left| \frac{S(j\omega)}{1 + G(j\omega)\, S(j\omega)} \right| = \left| \frac{1}{j\omega + G(j\omega)} \right| \tag{2}$$

The utility of the error response ratio is twofold: 1) With respect to system design, it places in evidence those frequency intervals of particularly high sensitivity of LOS error to the rate disturbance, thereby guiding equalization adjustments to the tracking loop; and 2) With respect to system analysis, the product of the square of the error response ratio, $E^2(\omega)$, times the spectrum of the load, $\Phi_{dd}(\omega)$ yields that part of the LOS error spectrum, $\Phi_{\epsilon\epsilon}(\omega)$, which is attributable to the load, the remainder being attributable to the remnant.

From equation (2), the effect of neglecting the sighthead bandwidth on the error response ratio may be assessed. Beyond the tracking loop crossover frequency, where the magnitude of the open-loop response, $[G(j\omega)\ S(j\omega)]$, is considerably less than one, the error response ratio is approximated by $\left|S(j\omega)\right| \leq \omega^{-1}$. Since the gain of the sighthead actually becomes less than ω^{-1} beyond its effective break frequency, 85.7 radians/second, the fact is evident that neglecting the bandwidth (that is, assuming an infinite bandwidth) results in a slight over estimate for $E(\omega)$. For example, at a frequency of 31.4 radians/second (5 Hz) the error response ratio would be .55 dB too high.

Next, consider some general characteristics of the gunner's load. The aspect of the load which taxes the capabilities of the gunner is not so much its magnitude (assuming adequate thumb controller authority) but rather its bandwidth. From the development in the Appendix, it is evident that the following factors can impart a rate of change or modulation to the load:

- The angular rate of the sighthead, particularly about the LOS
- Changes in range
- Changes in the magnitude or direction of the target velocity vector.

Factors which impart additional rates of change or modulation to the load are directly translatable into increased load bandwidth.

By way of elaboration on these factors, it will be recalled that the sighthead provides stabilization of the LOS with respect to aircraft rates perpendicular to the LOS; but, there is nothing to prevent the sighthead from rotating about the LOS. Hence, aircraft attitude changes about the LOS contribute very substantially to the modulation of the gunner's load. Furthermore, any translation of the aircraft will generally result in an inertial sighthead rate about an axis perpendicular to the LOS as well as a change in range.

Consider a typical engagement of the Cheyenne aircraft with a moving ground target. Suppose the engagement begins at a range of 2000 meters and terminates at a range of 1000 meters. If the aircraft velocity is 150 knots (77 meters/second) the encounter would last about 13 seconds. If the target velocity is 30 knots (15.4 meters/second) along a horizontal direction generally perpendicular to the LOS, the tracking control output rate which the gunner must supply in train will range from a minimum of about 7.7 milliradians/second to a maximum of 15.4 milliradians/second. The elevation rate will be very small. Barring any aircraft maneuvers, the mean or steady portion of the train rate would be about 11.6 milliradians/second and the varying portion would have a standard deviation of about 2.2 milliradians/second. An experienced gunner could supply the constant rate virtually without error, whereas the variable part of the rate would be acted upon by an overall response ratio, η, of about 0.15, so that the resulting standard deviation of tracking error in train would be in the vicinity of 0.33 milliradians. Suppose that during this pass the aircraft were to enter a coordinated 1.5g turn at the 1500 meter point. The overall effect of this maneuver would be to increase the bandwidth of the load and to shift a sizable fraction of it into the elevation axis. Whereas the elevation error in the previous case was very small, in this modified pass the elevation error would probably exceed that in train.

To summarize then, the magnitude and bandwidth of gunner's load will increase as a function of the relative velocity between aircraft and target and as a function of the angular rate of the aircraft about the LOS. If σ_d is increased while $\Phi_{dd}(0)$ is kept constant, thereby increasing the load bandwidth, then the overall error response ratio, $\eta = \sigma_\epsilon/\sigma_d$, will be caused to increase. As a convenient benchmark, an experienced tracker can readily achieve values of η between 0.17 and 0.20 when the bandwidth of the load is about 0.1 Hz.

THE GUNNER'S REMNANT

As shown in Figure 6, the gunner is represented by the combination of a linear response model, $G(j\omega)$, and an incoherent noise process or remnant, $n(t)$. In accordance with the human controller remnant model of Ref. 2, the remnant is regarded as entering the tracking loop at the input to the gunner. In other words, all the features of the gunner's response which produce remnant are represented in the effective observation noise, $n(t)$. The fact is evident that they could have been represented with equal validity in an effective actuation noise signal added to the gunner's output.

The spectral density of the remnant is conveniently defined in terms of the spectral density of the LOS error which is incoherent with the tracking loop forcing function, d(t). Toward this end, the following definitions are made: let $\Phi_{nn}(\omega)$ denote the spectral density of the remnant itself (i.e., the open-loop remnant spectrum); let $\Phi_{nn_c}(\omega)$ denote the spectral density of the incoherent part of the LOS error, or that part of the LOS error which is not linearly correlated with the load (i.e., the closed-loop remnant spectrum); and let $T(\omega)$ denote the closed-loop response ratio of the tracking loop. In terms of the elements of the tracking loop, $T(\omega)$ is given by

$$T(\omega) = \left| \frac{G(j\omega)\, S(j\omega)}{1 + G(j\omega)\, S(j\omega)} \right| = \left| G(j\omega) \right| \; E(\omega) \tag{3}$$

By means of the closed-loop response ratio, the open and closed-loop remnant spectra may be conveniently related as follows:

$$\Phi_{nn_c}(\omega) = T^2(\omega)\, \Phi_{nn}(\omega) \tag{4}$$

But since the closed-loop remnant is, in fact, the spectrum of the incoherent LOS error, it is equivalent to the difference between the total LOS error spectrum, $\Phi_{\epsilon\epsilon}(\omega)$, and the coherent LOS error spectrum, $[E^2(\omega)\, \Phi_{dd}(\omega)]$. That is,

$$\Phi_{nn_c}(\omega) = \Phi_{\epsilon\epsilon}(\omega) - E^2(\omega)\, \Phi_{dd}(\omega) \tag{5}$$

A quantity which is of great value in assessing the relative importance of the remnant in a given control task is the ratio of coherent LOS error variance divided by the total LOS error variance. This ratio shall be referred to as the coherence index and shall be denoted by ξ, its mathematical definition being as follows:

$$\xi^2 = \frac{\int_0^\infty E^2(\omega)\Phi_{dd}(\omega)\, d\omega}{\sigma_\epsilon^2} \tag{6}$$

Using equation (5) to solve for the integrand, $[E^2(\omega)\, \Phi_{dd}(\omega)]$, and denoting the variance of the closed-loop remnant by $\sigma_{n_c}^2$,

$$\xi^2 = 1 - \frac{\int_0^\infty \Phi_{nn_c}(\omega)\, d\omega}{\sigma_\epsilon^2} = 1 - \sigma_{n_c}^2 / \sigma_\epsilon^2 \tag{7}$$

The fact is evident that ξ^2 lies between 0 and 1, a value near 0 indicating that the LOS error is predominantly the result of remnant, and a value near 1 indicating it is predominantly the result of a linear operation on the load, d(t). A typical value for ξ^2 appears to be about 2/3. This means that the total LOS error variance is about 3/2 times the coherent LOS error variance, $\int_o^\infty E^2(\omega)\, \Phi_{dd}(\omega)\, d\omega$.

The point should be made that the ratio ξ^2 defined above is akin to the "relative remnant", ρ_a^2, defined in Ref. 3 (p. 120) as the coherent operator output variance divided by the total operator output variance.

According to Ref. 2, if the remnant spectrum, $\Phi_{nn}(\omega)$, is normalized by the LOS error variance, the resulting spectrum is essentially invariant with respect to changes in input bandwidth, input injection point, and controlled element dynamics. This normalized remnant spectrum shall be denoted $\Phi^*_{nn}(\omega)$ (i.e., the normalized open-loop spectrum).

$$\Phi^*_{nn}(\omega) = \Phi_{nn}(\omega)/\sigma_\epsilon^2 \tag{8}$$

The normalized closed-loop spectrum is denoted by $\Phi^*_{nn_c}(\omega)$, and the corresponding variance by $\sigma^{*2}_{n_c}$.

$$\Phi^*_{nn_c}(\omega) = T^2(\omega)\, \Phi^*_{nn}(\omega) \tag{9}$$

$$\sigma^{*2}_{n_c} = \int_o^\infty \Phi^*_{nn_c}(\omega)\, d\omega \tag{10}$$

In the terms of $\sigma^{*2}_{n_c}$, the coherence ratio, ξ^2, reduces to the following simple expression:

$$\xi^2 = 1 - \sigma^{*2}_{n_c} \tag{11}$$

The neglect of the bandwidth of the sighthead will have a slight effect upon the remnant spectrum. Since the error response ratio is slightly too high, the fact is evident from equation (5) that the remnant spectrum computed from a given sample of tracking data must be correspondingly low. Actually, the effect of this assumption is best expressed in terms of the coherence ratio, ξ^2. The computed value of ξ^2 is slightly too high since some of the total LOS error variance which was actually produced by remnant is in the end attributed to the operation of $E^2(\omega)$ on $\Phi_{dd}(\omega)$.

Section II

THE PROBLEM OF TRACKING ERROR PREDICTION

The purpose of the foregoing was largely to provide a background for this statement: Reduced to its simplest terms, the problem of predicting tracking error is a three-step process, the first step involving the computation of the gunner's load for a given engagement, the second amounting to the determination of the tracking loop error response ratio, and the third involving the operation on the load by the error response ratio to obtain the error.

The first and third steps of the process are straightforward. It is the second step, the determination of the error response ratio, which presents the real challenge in the overall prediction problem. Furthermore, the sighthead dynamic characteristics are well known so that selection of the gunner's describing function is tantamount to determining the error response ratio.

The remainder of this paper concerns the selection of the describing function model for the gunner from published data, and comparing that result with actual measurements on the SGS.

Section III

THE SELECTION OF A GUNNER DESCRIBING FUNCTION MODEL

Two approaches to selecting a describing function model for the gunner will be discussed. The first approach is through application of the pilot describing function results published in Ref. 3, and the second is through the use of experimental data obtained in the GE/ACD Dynamic Simulation Laboratory.

Note that the first approach is that of a knowledgeable system analyst who is faced with the problem of predicting tracking accuracy for a sighting station such as the SGS before the equipment has been built and tested. The second is that of the same analyst after the equipment has been built and some system performance data have been obtained. Then the problem is basically one of generalizing on the available test data to predict system performance in particular operational (that is, field) environments. The question is, how well does the describing function based on published tracking results compare with that extracted from experimental tracking results, insofar as LOS error prediction is concerned?

In the case of both these describing functions, a certain amount of intuition was exercised with regard to the very low frequency response of an SGS operator. While referring to the tracking loop block diagram of Figure 6, suppose the rate disturbance or load, d, is a constant value. Intuitively it is reasonable to expect that after a short time (in the order of 1 second) the gunner will have not only developed the proper thumb force to nullify this constant disturbance rate, but he will have also eliminated virtually all of the LOS error. In other words, not only will the LOS error rate be driven to zero, but also the LOS error angle*. To embody this low frequency response characteristic in a linear describing function model requires a structure which is essentially that given below:

$$G(j\omega) = \frac{K\left(1 + \frac{j\omega}{\omega_L}\right) e^{-j\omega \tau_e}}{j\omega} \qquad (12)$$

*In actual fact, when the rate disturbance forcing the tracking system contains a mean rate in addition to a random component, experimental results indicate that the human subject does indeed perform according to the above intuitive expectation.

That is, the operator supplies what amounts to an integrator-lead equalization in the very low frequency range. When this is combined with the integral response of the sighthead, the velocity error coefficient of the tracking loop becomes infinite.

In the case of the describing function based on the results published in Ref. 3, the imposition of this structural requirement results in an alternate form of extended crossover model. This same low frequency structure is imposed on the model extracted from the experimental data which is described below.

The experimental describing function data suggest that the behavior of the gunner in the elevation and train axes of control is somewhat different. For this reason, the results obtained in the two control axes will be displayed separately.

The following symbols will be employed in the ensuing discussion:

$G_C(j\omega)$ = The gunner's describing function corresponding to the extended crossover model

$L_C(j\omega)$ = $G_C(j\omega)/j\omega$

= The extended crossover model of the open-loop response ratio

$E_C(\omega)$ = The error response ratio corresponding to the extended crossover model

$G_E(j\omega)$ = The gunner's describing function in the elevation axis of control as obtained from SGS tracking data

$L_E(j\omega)$ = $G_E(j\omega)/j\omega$

= The experimental open-loop response ratio in the elevation axis

$E_E(\omega)$ = The experimental error response ratio in the elevation axis

$G_T(j\omega)$ = The gunner's describing function in the train axis of control as obtained from SGS tracking data

$L_T(j\omega)$ = $G_T(j\omega)/j\omega$

= The experimental open-loop response ratio in the train axis

$E_T(\omega)$ = The experimental error response ratio in the train axis.

DESCRIBING FUNCTION SELECTION BASED ON PUBLISHED DATA

The model $G_C(j\omega)$ is based upon the extended crossover model for $Y_C = K_C(j\omega)^{-1}$ set forth on page 153 of Ref. 3. Specifically, equation (92) on this page states that the phase effect of very low frequency lag-lead equalization by the operator may be modeled in the vicinity of crossover by a slight reduction in the value of the effective transport lag, as compared with that used in a straight crossover model, combined with an addition to the transport lag of (α/ω^2). However, this adjustment does not impart the gain effect of such low frequency equalization, which is required for a describing function model suited to the prediction of SGS operator tracking behavior.

Accordingly, the assumption is made that the (α/ω^2) contribution to the transport lag is for all practical purposes equivalent to a lag at zero frequency plus a low frequency lead. Hence, the lead break-frequency is sought for which

$$\tan^{-1}\left(\frac{\omega}{\omega_L}\right) - \frac{\pi}{2} = -\frac{\alpha}{\omega} \tag{13}$$

Provided that ω is significantly higher than ω_L, the phase shift of the lead may be approximated by $\left(\frac{\pi}{2} - \frac{\omega_L}{\omega}\right)$; thus, the result obtains that ω_L should be made equivalent to α. Modifying equation (92) of Ref. 3 in accordance with this result gives

$$Y_p Y_c = \frac{\alpha\overline{\omega}_c \left(1 + \frac{j\omega}{\alpha}\right) e^{-j\omega\tau_e}}{(j\omega)^2} \tag{14}$$

Table X of Ref. 3 gives $\overline{\omega}_c$ as 4.75 radians/second and α and τ_e are given as functions of input bandwidth. In the case of the SGS tracking loop, the low-pass rate disturbance is additionally filtered by the integral controlled element so that the effective reference input bandwidth is zero for all practical purposes. This indicates the use of $\tau_e = \overline{\tau}_o = 0.33$ second. The selection of an appropriate value for α is not as clear-cut due to the non-monotonic relationship between α and ω_i, which makes extrapolation back to $\omega_i = 0$ somewhat arbitrary. Accordingly, the average of the three tabulated α-values is chosen; namely, $\alpha = 0.15$ radian/second.

Thus, using the nomenclature established above, an extended crossover model appropriate for the prediction of SGS tracking accuracy is:

$$G_C(j\omega) = \frac{.713\,(1 + \frac{j\omega}{.15})\, e^{-j\,.33\omega}}{j\omega} \tag{15}$$

Replacing the transport lag by a Padé approximation, the following equivalent model is obtained:

$$G_C(j\omega) = \frac{.713\,(1 + \frac{j\omega}{.15})\,(1 - \frac{j\omega}{6.06})}{j\omega\,(1 + \frac{j\omega}{6.06})} \tag{16}$$

The corresponding open-loop function is

$$L_C(j\omega) = G_C(j\omega)/j\omega \tag{17}$$

and the error response ratio is

$$E_C(\omega) = \left| \frac{1.40\; j\omega\,(1 + \frac{j\omega}{6.06})}{(1 + \frac{j\omega}{.155})\left[1 + .218\left(\frac{j\omega}{5.28}\right) + \left(\frac{j\omega}{5.28}\right)^2\right]} \right| \tag{18}$$

DESCRIBING FUNCTION SELECTION BASED ON EXPERIMENTAL DATA

Commencing in the latter part of 1968, a serious experimental effort was initiated to obtain an accurate describing function model of an SGS operator under a variety of forcing function conditions, the experimental phase of this project being conducted in the GE/ACD Dynamic Simulation Laboratory.

A single subject acted as the SGS operator for most of these experiments. Although his background did not include any experience as a pilot, he did have considerable experience as a subject in laboratory tracking experiments. In particular, he had accumulated about 50 hours of tracking time on the SGS in the Dynamic Simulation Laboratory and, prior to that, about 20 hours experience in early tracking experiments related to the SGS tracking control design.

Figure 7 shows an overall view of the laboratory area. In the rear is seen the base motion simulator on which is mounted an SGS. The simulator produces angular motions about two axes to model aircraft yaw and roll attitude variations. The simulator console is seen in the left-foreground of the picture. In the right-foreground are seen the following pieces of equipment:

- The SGS Test Panel

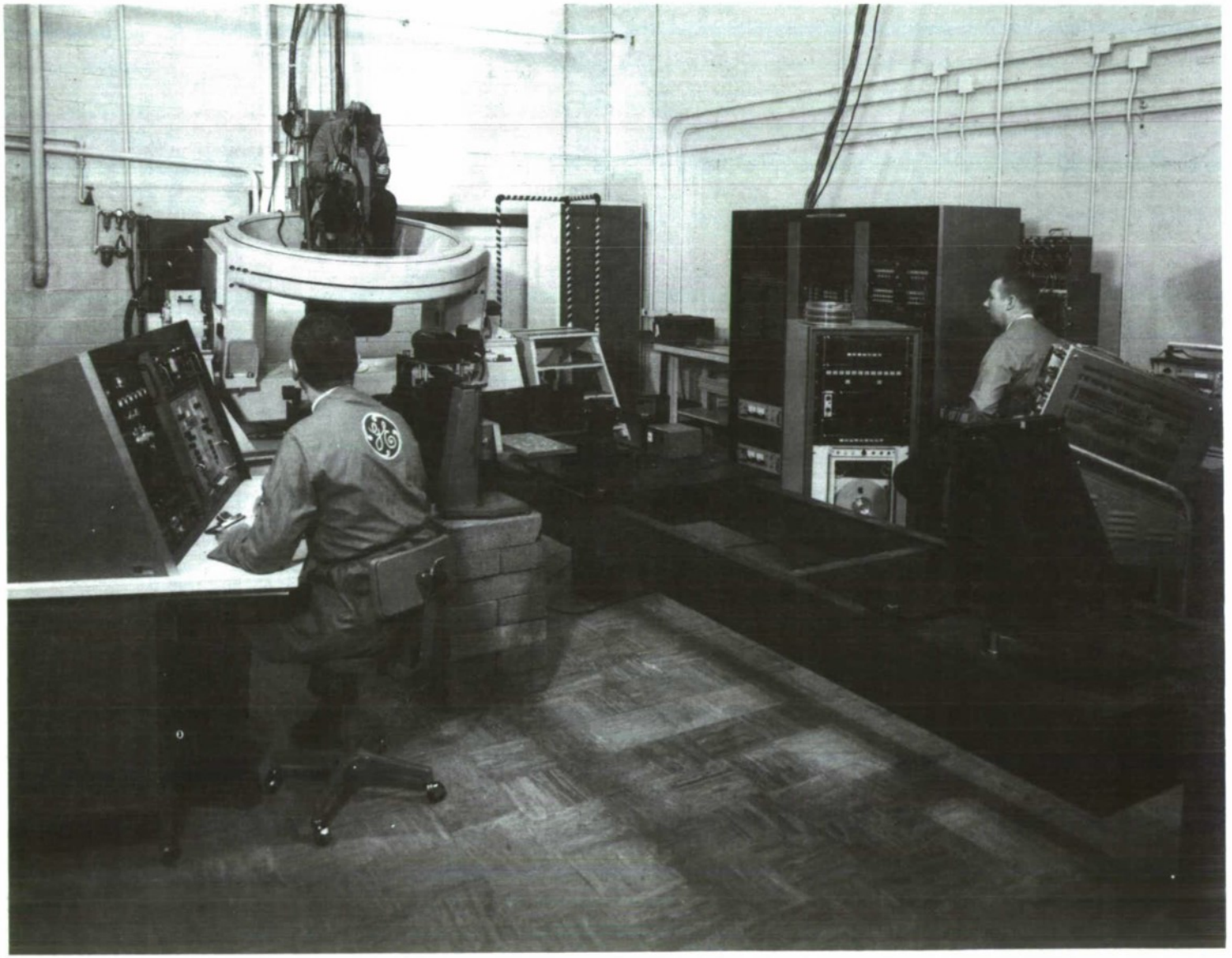


Figure 7. An Overall View of the Dynamic Simulation Laboratory

- Analog computers used to generate the tracking-loop disturbance and filter the output data

- A datalogger consisting of a multiplexing analog/digital converter and a digital magnetic tape recorder.

In the center-foreground is a fixture on which is mounted a 12-inch collimating mirror, the purpose of which was to project a stationary, collimated target image for the SGS operator to track.

Between the 12-inch collimator fixture and the base motion simulator is a fixture on which is mounted a moving target projector. This apparatus was not used during these particular experiments.

The outer and horizontal gimbal of the base motion simulator is considered to represent the roll axis of the aircraft (accordingly, the gunner is operating the SGS about a nominal train angle of 90 degrees with respect to the aircraft centerline). The inner, and nominally vertical, gimbal of the simulator is considered to represent the aircraft's yaw axis.

The roll excitation of the base motion simulator was a bandlimited noise source having a standard deviation of about 10 degrees and a bandwidth of about 0.2 radian/second. A similar but independent source excited the yaw motion, its standard deviation being about 30 degrees and its bandwidth, also 0.2 radian/second.

Twenty-five tracking runs were made with the standard deviation of the gunner's load, σ_d, being varied from about .5 milliradian/second to about 50 milliradians/second, the bandwidth being held constant at about .1 Hz. The duration of each run was 33.8 seconds.

Figure 8 shows the 25-run average of the gunner's load spectra applied in the elevation and train axes of control. The high frequency shelf was employed to place in evidence the characteristics of the gunner beyond the open-loop crossover frequency of about .75 Hz or 4.7 radians/second.

Since the target image in these experiments was stationary, and since target motion was effectively simulated by the load, d, any rate command which reached the sighthead servoes corresponded to an LOS error rate, labeled $\dot{\epsilon}$ in Figure 6. The signals which were recorded on the datalogger were the load, d, and the LOS error rate, $\dot{\epsilon}$, in each axis.

Actually, the error rate was sampled twice, once immediately preceding and once immediately following the sampling of the load. Thus, in the course of processing the data, the two error rate samples could be averaged to yield an approximation to the value which coincided with the sample of the load. The coincidental value of the tracking control output, h, was computed from the values of d and $\dot{\epsilon}$. In addition to this, the error rate time history was numerically integrated to obtain the LOS error angle, ϵ.

Well established spectral and cross-spectral analysis procedures were applied to the time histories for d, h, $\dot{\epsilon}$, and ϵ to obtain the open-loop function, the error response ratio and the remnant spectrum in each axis of control (see Ref. 4, pages 31-43 for background). With Φ_{xx} denoting the spectral density of a generic quantity x, Φ_{xy}, the cross-spectral density of x with respect to y, and Ave (Φ) the average or mean spectral value over 25 runs, the following relationships describe the procedure which was employed:

$$L(j\omega) = \text{Ave}\,(\Phi_{dh})/\text{Ave}\,(\Phi_{d\dot{\epsilon}}) \tag{19}$$

$$E(\omega) = \left|\text{Ave}\,(\Phi_{d\epsilon})/\text{Ave}\,(\Phi_{dd})\right| \tag{20}$$

$$T(\omega) = \left|\text{Ave}\,(\Phi_{dh})/\text{Ave}\,(\Phi_{dd})\right| \tag{21}$$

$$\Phi^*_{nn} = \frac{\text{Ave}\,(\Phi_{\epsilon\epsilon}) - E^2(\omega)\,\text{Ave}\,(\Phi_{dd})}{\sigma^2_{\epsilon}\;T^2(\omega)} \tag{22}$$

By way of clarification of the averaging operator used in equations (19-22), if $\Phi_i(\omega)$ denotes the spectral value measured at the frequency ω in the i^{th} run, then:

$$\text{Ave}\,[\Phi(\omega)] = \frac{1}{25}\sum_{i=1}^{25}\Phi_i(\omega) \tag{23}$$

RESULTS

In Figures 9-11 are plotted the gain and phase of the open-loop function, $L(j\omega)$, and the error response ratio, $E(\omega)$, for the elevation and train control axes. The experimental data obtained by means of the procedure described above are plotted as large dots, and the functions corresponding to the extended crossover model are plotted as dotted lines. The functions corresponding to a gunner model which was fitted to the experimental data are

plotted as solid lines. This fit was accomplished through a parameter optimization routine which operated to minimize a performance index including the following items:

- The sum-square deviation of the gain of the calculated open-loop function from the measured values in decibels
- The sum-square deviation of the phase of the calculated open-loop function from the measured values in degrees
- The sum-square deviation of the calculated error response ratio from the measured values in decibels
- A penalty function reflecting the violation of certain minimum stability requirements (specifically, that the phase margin be at least 10 degrees, and the gain margin be at least 3 decibels).

The experimental gunner model functions corresponding to the extended crossover model functions given in equations (16-18) are set forth in equations (24-29).

$$G_E(j\omega) = \frac{.727\left(1+\frac{j\omega}{.0767}\right)\left(1+\frac{j\omega}{1.22}\right)\left(1-\frac{j\omega}{10.8}\right)}{j\omega\left(1+\frac{j\omega}{.546}\right)\left(1+\frac{j\omega}{10.8}\right)} \tag{24}$$

$$L_E(j\omega) = G_E(j\omega)/j\omega \tag{25}$$

$$E_E(\omega) = \frac{j\omega\left(1+\frac{j\omega}{.546}\right)\left(1+\frac{j\omega}{10.8}\right)}{.727\left(1+\frac{j\omega}{.0773}\right)\left(1+\frac{j\omega}{1.48}\right)\left[1+.906\left(\frac{j\omega}{6.12}\right)+\left(\frac{j\omega}{6.12}\right)^2\right]} \tag{26}$$

$$G_T(j\omega) = \frac{4.01\left(1+\frac{j\omega}{.428}\right)\left(1+\frac{j\omega}{.706}\right)\left(1-\frac{j\omega}{9.43}\right)}{j\omega\left(1+\frac{j\omega}{.359}\right)\left(1+\frac{j\omega}{9.43}\right)} \tag{27}$$

$$L_T(j\omega) = G_T(j\omega)/j\omega \tag{28}$$

$$E_T(\omega) = \frac{j\omega\left(1+\frac{j\omega}{.359}\right)\left(1+\frac{j\omega}{9.43}\right)}{4.01\left(1+\frac{j\omega}{.421}\right)\left(1+\frac{j\omega}{.854}\right)\left[1+.610\left(\frac{j\omega}{6.15}\right)+\left(\frac{j\omega}{6.15}\right)^2\right]} \tag{29}$$

In Table 1 are listed those parameters which are especially useful in comparing the extended crossover model with the experimentally based gunner model.

The describing function characteristics which best reflect the tracking behavior of the gunner are the open-loop crossover frequency and the effective transport lag. It is noteworthy that the experimental results are in remarkably good agreement with the extended crossover model in regard to the crossover frequency. The average of the two experimental values in Table 1 is 4.6 radians/second, whereas the value corresponding to the extended crossover model is 4.8 radians/second. There is a considerable discrepancy between the values for the effective transport lag, however. The average experimental value is about .20 second, whereas that for the extended crossover model is .33 second. This difference is reflected in the larger experimental phase margins, which average to about 31 degrees, as compared with a phase margin of only 13 degrees for the extended crossover model. It is also reflected in the greater closed-loop peaking evidenced in the error response ratio, E_C, as compared with E_E and E_T in Figure 11. In fact, E_C exceeds E_E by a maximum amount of about 8 dB in the vicinity of the crossover frequency, and E_T by about 6 dB.

The describing function characteristics which are of the greatest concern in the present context are those which affect the prediction of tracking error for a given gunner's load spectrum (see Figure 6). These characteristics are reflected in an articulate fashion by the error response ratios plotted in Figure 11, and in a gross fashion by the overall error response ratios given in the first row of Table 1.

The extended crossover model gives rise to an overall error response ratio of about .28. This is 47 percent higher than the average of the two experimental error response ratios, which is about .19. The difference is due primarily to the fact that the error response ratio E_C exceeds E_E and E_T in the low frequency range from .03 Hz to .3 Hz (.19 radian/second to 1.9 radians/second). It is due secondarily to the greater amount of closed-loop peaking in E_C, which was discussed above. The fact that E_E and E_T fall below E_C in the low frequency range is the direct result of the gunner's employing a higher gain in this part of the spectrum than the extended crossover model predicts. To accomplish this and yet maintain the same crossover frequency, he evidently introduces a lag-lead compensation term, as reflected in equations (24) and (27) for G_E and G_T, respectively.

TABLE 1

COMPARISON OF SIGNIFICANT DESCRIBING FUNCTION PARAMETERS

Parameter	Extended Crossover Model	Experimental Gunner Model	
		Elevation Axis	Train Axis
Overall Error Response Ratio	.28	.20	.18
Effective Transport Lag	.33 Sec.	.18 Sec.	.21 Sec.
Open-Loop Crossover Freq.	.76 Hz or 4.8 Rad/Sec	.70 Hz or 4.4 Rad/Sec	.77 Hz or 4.8 Rad/Sec
Phase Margin	13^{o}	36^{o}	27^{o}

Significant experimental results were also obtained for the normalized open-loop remnant spectrum, Φ^*_{nn}, defined in equation (8). These results are plotted in Figures 12a and 12b for the elevation and train axes of control. For purposes of comparison, the human controller remnant model of Ref. 2 is plotted as a dotted line, after adjustment for the change in frequency scale from radians/second to Hz and the change in the spectral density units from power per radian/second to power per Hz.

The agreement between the experimental and published remnant spectra is regarded as good, especially for frequencies beyond crossover (that is, beyond about .75 Hz or 4.7 radians/second). The fact that the experimental remnant spectrum exceeds the model below crossover is probably significant. Two explanations for this difference appear plausible:

- Since the experimental data were obtained in the context of a two-axis control task, the forcing functions in the two axes being mutual independent, any cross-coupling in the gunner's response would ultimately be interpreted as an addition to the remnant.

- Another feature of the experimental environment was the random base motion input (particularly that in roll, the full magnitude of which was experienced by the gunner). Any propagation of this motion into the tracking control would further augment the remnant since the base motion forcing functions and those disturbing the tracking loop were all mutually independent.

Since the logical extension of the second explanation would be that the remnant augmentation due to base motion should be substantially more pronounced in the elevation axis than in train, and since the results plotted in Figure 12 do not bear this out, the conclusion is drawn that the bulk of the excess remnant is due to cross-coupling in the gunner's response.

An overall measure of the importance of the remnant in a given control task is the coherence index, ξ^2, defined in equations (7) and (11). The values for ξ^2 which correspond to the remnant spectra plotted in Figure 12 are .70 in elevation and .64 in train. These average to a value of .67 which means that in the SGS tracking task about 2/3 of the LOS error variance is predictable by a linear operation upon the gunner's load spectrum whereas the remaining 1/3 is the result of remnant. In other words, the LOS error variance which is obtained through the use of a linear gunner response model should be multiplied by a factor of about 3/2 to estimate the total LOS error variance, including remnant effects.

CONCLUSIONS

For the gunner's compensatory tracking task (pure rate controlled element and integrated low-pass input spectrum) it was found that the experimentally measured gunner describing function agrees remarkably well with an alternate form of the extended crossover model in regard to open-loop crossover frequency. The experimental results indicate that below crossover the gunner introduces lag-lead compensation terms which permit an elevated low frequency gain, and hence, superior error attenuation performance. Although this feature of the gunner's behavior constitutes a deviation with respect to the extended crossover model, it is certainly allowable within the context of more complicated model forms. In view of the dominant motivation of the gunner to minimize tracking error, his apparent use of lag-lead compensation below crossover is not unexpected.

The experimental results manifest a substantially higher phase margin than that predicted by means of the extended crossover model. This difference may be stated in the following manner: although the SGS gunner was responding to a very low bandwidth input, his effective transport lag was as if he were responding to a rather high bandwidth input. His behavior results in a substantial reduction in closed-loop peaking which, in turn, results in superior error attenuation performance. The tentative conclusion to be drawn from this is that the adjustment rules which describe the variation of τ_e with ω_i for the average experienced pilot subject do not necessarily apply in the case of an SGS operator. The latter is highly motivated to minimize tracking error and is relatively unconcerned over the demands he makes on his controlled element to achieve that objective.

The measured normalized open-loop remnant is in agreement with a recently published model for controller remnant for frequencies above crossover and exceeds the model by a significant amount below crossover, especially at very low frequencies. The most plausible explanation for this difference appears to be cross-coupling in the gunner's response in the two-axis SGS tracking task, which was characterized by mutually independent forcing functions in elevation and train.

The accuracy of an operator describing function for application to SGS tracking error prediction is reflected by the overall error response ratio. The value for this parameter corresponding to the extended crossover model proved to be almost 50 percent higher than the average of the measured values.

Accordingly, the conclusion is drawn that a direct application of the extended crossover model, together with its adjustment rules, leads to a rather conservative estimate of SGS tracking performance. In the case of the SGS, the need for an accurate prediction of tracking performance was such as to justify the in situ measurement of a gunner's describing function model.

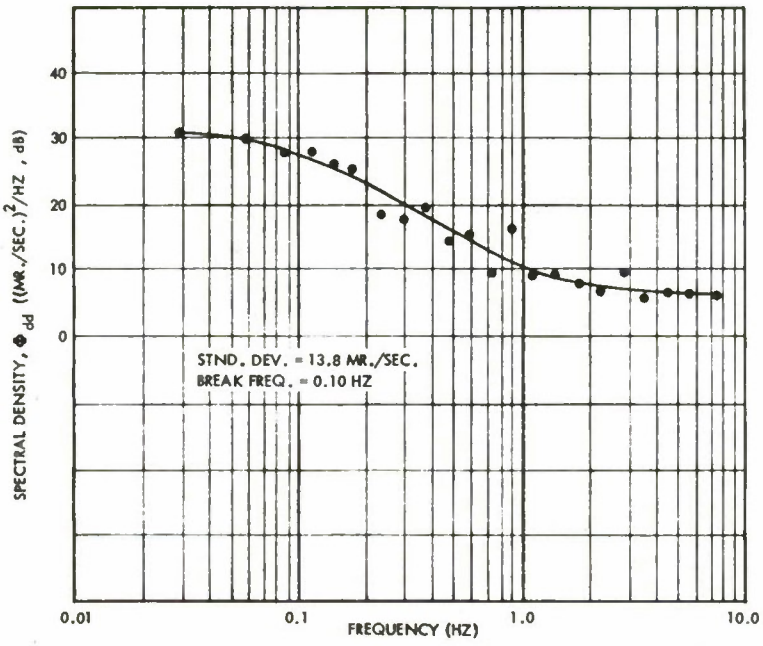


Figure 8a. Tracking Loop Disturbance - Elevation Axis

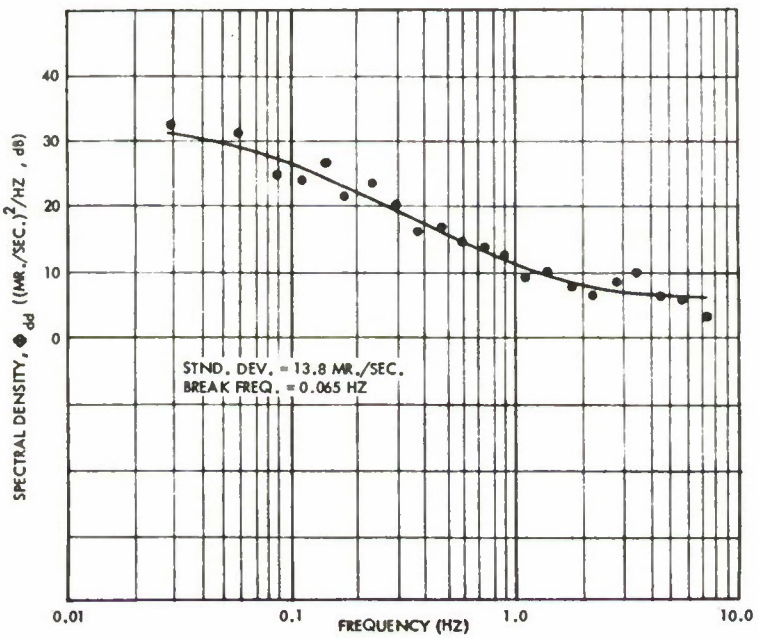


Figure 8b. Tracking Loop Disturbance - Train Axis

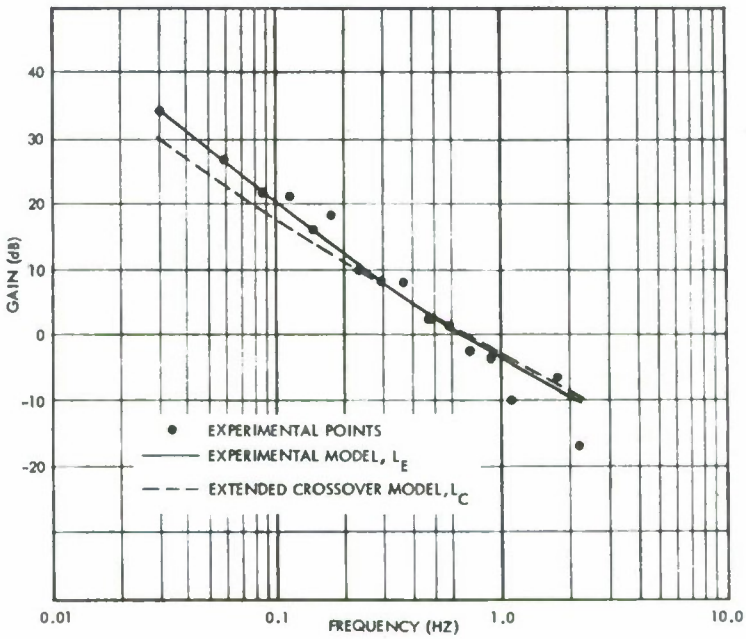


Figure 9a. Open-Loop Function - Elevation Axis

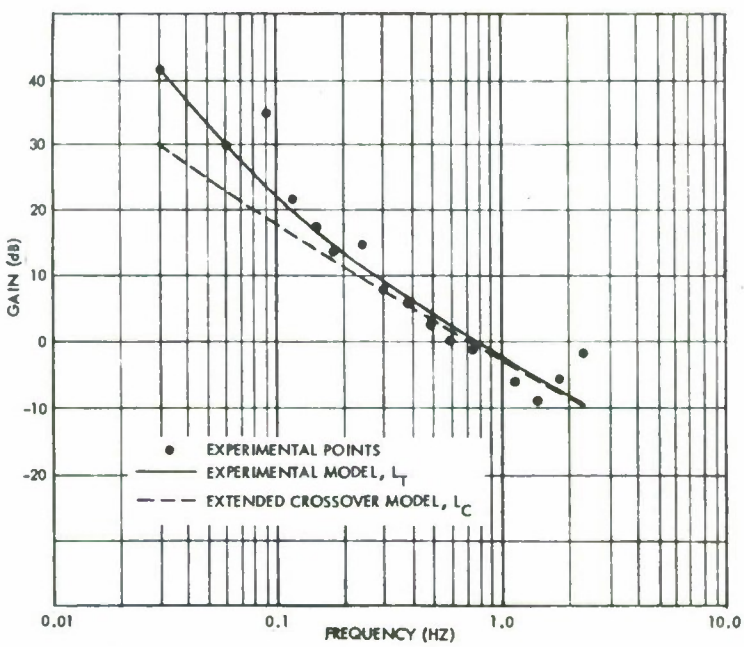


Figure 9b. Open-Loop Function - Train Axis

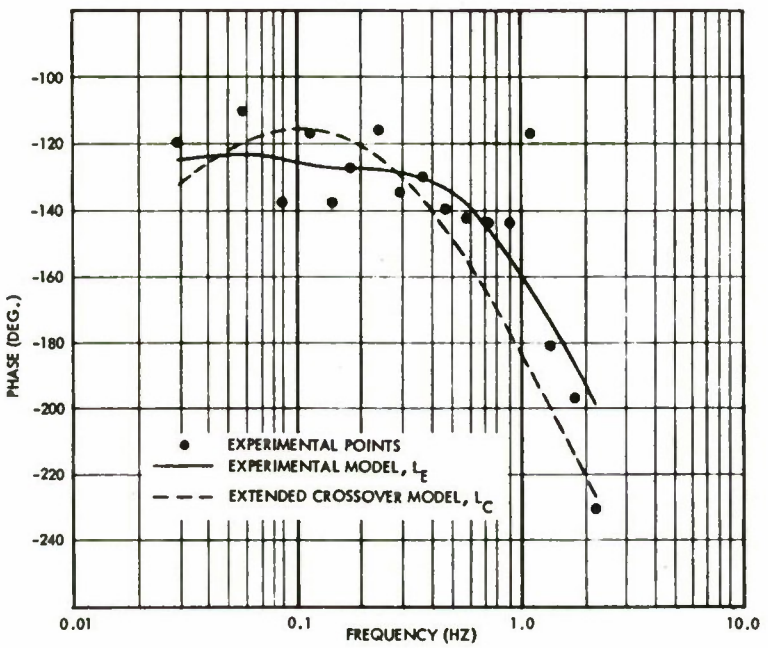


Figure 10a. Open-Loop Function - Elevation Axis

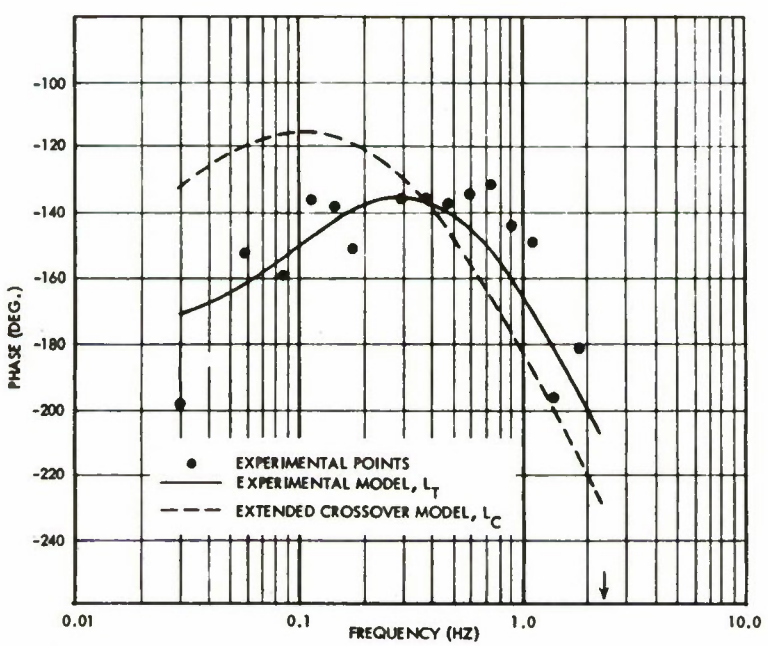


Figure 10b. Open-Loop Function - Train Axis

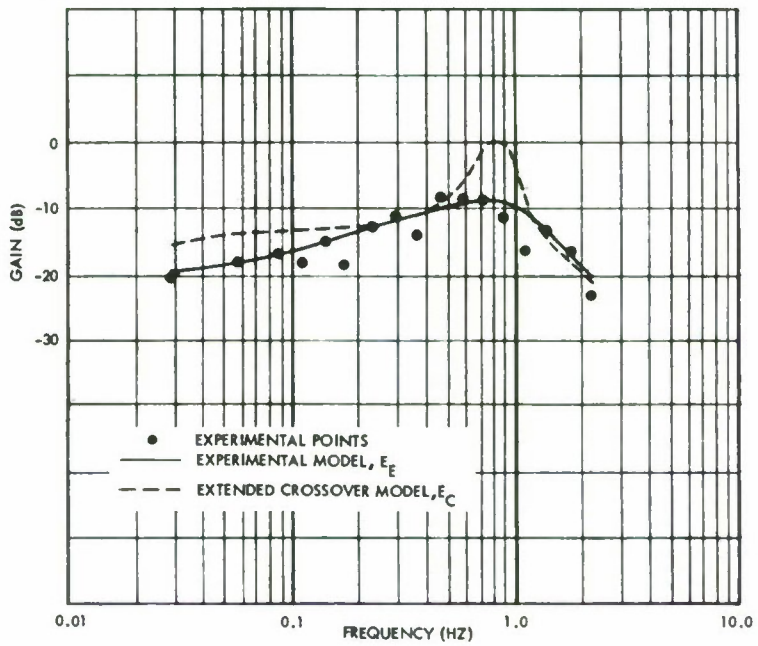


Figure 11a. Error Response Ratio - Elevation Axis

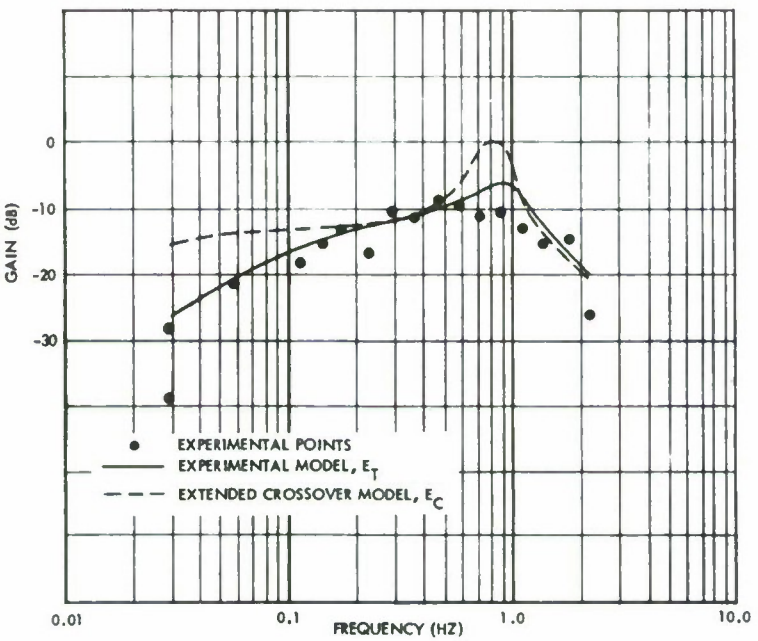


Figure 11b. Error Response Ratio - Train Axis

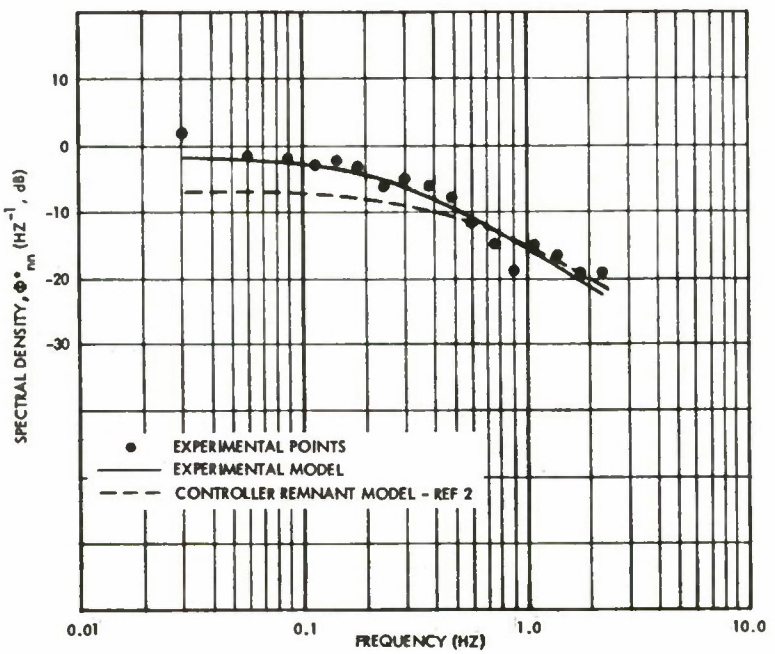


Figure 12a. Normalized Open-Loop Remnant - Elevation Axis

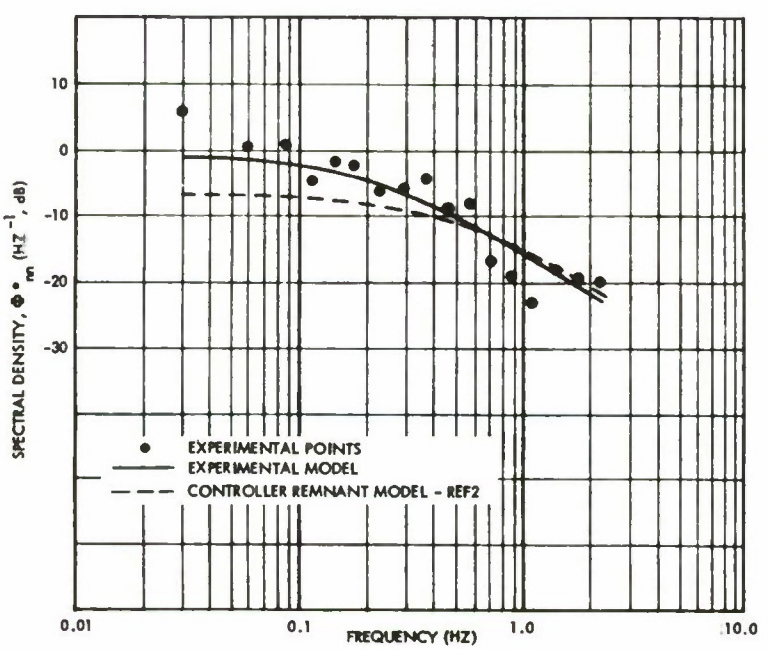


Figure 12b. Normalized Open-Loop Remnant - Train Axis

APPENDIX

DETERMINISTIC EQUATIONS FOR THE GUNNER'S LOAD

The kinematic relations describing an engagement of the Cheyenne aircraft with a moving ground target are used to develop deterministic equations for the gunner's load. The engagement geometry is shown in Figure 5.

Letting $\overline{V}_A$ denote the velocity of the aircraft, $\overline{V}_T$ that of the target, R the range, and $\overline{s}$ the LOS position, the angular rate at which the LOS must be slewed is:

$$\overline{\omega} = \frac{1}{R} \left[(\overline{V}_A - \overline{V}_T) \times \overline{s} \right] \tag{30}$$

It is the function of the LMC to provide the rate

$$\overline{\omega} = \frac{1}{R} (\overline{V}_A \times \overline{s}) \tag{31}$$

so that the gunner is left with the rate

$$\overline{\omega}_H = \frac{-1}{R} (\overline{V}_T \times \overline{s}) + \delta\overline{\omega}_c \tag{32}$$

where $\delta\overline{\omega}_c$ denotes the LMC error.

Let $\overline{\ell}$ be the axis perpendicular to $\overline{s}$ about which elevation motions of the LOS occur, and let $\overline{m}$ be a third axis perpendicular to both $\overline{s}$ and $\overline{\ell}$ such that $\overline{s}$, $\overline{\ell}$, and $\overline{m}$ form a right-handed set. For visualization, $\overline{\ell}$ and $\overline{m}$ correspond to the horizontal and vertical crosshairs of the gunner's reticle display, and are therefore the elevation and train axes insofar as the gunner is concerned. Accordingly, the projection of $\overline{\omega}_H$ along $\overline{\ell}$ constitudes the load in elevation and that along $\overline{m}$ constitutes the load in train. These projections along $\overline{\ell}$ and $\overline{m}$ are designated respectively d_ℓ and d_m:

$$d_\ell = \overline{\omega}_H \cdot \overline{\ell} = \frac{-1}{R} (\overline{V}_T \times \overline{s}) \cdot \overline{\ell} + \delta\overline{\omega}_c \cdot \overline{\ell} \tag{33}$$

$$d_m = \overline{\omega}_H \cdot \overline{m} = \frac{-1}{R} (\overline{V}_T \times \overline{s}) \cdot \overline{m} + \delta\overline{\omega}_c \cdot \overline{m} \tag{34}$$

These two equations may be readily simplified to the following:

$$d_{\ell} = \frac{-\overline{V}_T \cdot \overline{m}}{R} + \delta\overline{\omega}_c \cdot \overline{\ell} \tag{35}$$

$$d_m = \frac{\overline{V}_T \cdot \overline{\ell}}{R} + \delta\overline{\omega}_c \cdot \overline{m} \tag{36}$$

Equations (35) and (36) show that the velocity of the target, $\overline{V}_T$, the range R, and the LMC error, $\delta\overline{\omega}_c$, all obviously contribute to the gunner's load in elevation and train. What is less obvious, however, is the modulation effect which occurs because of rotation of the sighthead. This effect arises in the dot products because of the changing angles between vectors.

REFERENCES

1. Allen, R. W., 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.

2. Levison, W. H., S. Baron, D. L. Kleinman, A Model for Human Controller Remnant, IEEE Trans., Vol. MMS-10, No. 4, December 1969.

3. McRuer, D. T., D. Graham, E. Krendel, W. Reisener, Jr., Human Pilot Dynamics in Compensatory Systems, AFFDL-TR-65-15, July 1965.

4. McRuer, D. T., E. Krendel, Dynamic Response of Human Operators, WADC TR-56-524, October 1957.

MAN/MACHINE FUNCTION ALLOCATION OF FLIGHT CONTROL

By

O. H. Lindquist
Systems and Research Division
Honeywell, Inc.

ABSTRACT

An acceptable man/machine function allocation for flight crew stations in aircraft requires answers to the following questions:

1. How much of the flight crew's time must be occupied by the flight control function for each mission segment?

2. How much flight crew time can be freed from flight control functions when additional automatic augmentation is introduced?

3. What quantitative description of vehicle flight control performance can be used to control work load and meet mission performance requirements?

Handling qualitites specifications and published information provide minimum assistance in resolving the problems.

This paper contains the following items:

- A definition of the problem
- A discussion of handling quality relations to mission performance
- Approach as to the problem

INTRODUCTION

Military aircraft have been found useful and necessary long after their introduction into service. Pilots and crews make surprisingly good use of equipment delivered 10 to 20 years ago. They devise new tactics, new procedures and generally use eyeballs, muscles and Kentucky windage to "make-do." In contrast, many automated systems have proven disappointing or useless in operation. They have failed for reasons categorized in terms

such as ECM, cost, reliability, maintainability; generally we have "effective" solutions to the wrong set of problems. The present situation of not being able to pre-plan and pre-program all alternates of the future will continue to plague all canned tactical solutions, however effective they are in meeting today's "scenarios."

The system design of flight vehicles of the future must consider the implication of the situation just described. The initial man/machine function allocation and consequent design of flight crew stations determine to a large degree whether the system will be useful in meeting new operational problems. The necessary condition in assuring that the crew will have a means to handle the undefined future problems is the availability of crew capability to apply to the problem. The system must be designed to permit the crew the opportunity for information processing, decision making and execution of function. Since in this case, these are unknown, the primary reservation of such capability must be in terms of available operator time.

A DEFINITION OF THE PROBLEM

For a given flight control alternate which might be recommended, how much of the flight crew's time is occupied in obtaining required system performance for each mission segment? The system design of the avionics for aircraft requires allocation of functions to the flight crew and his avionic equipment. The allocation is bounded by inherent vehicle characteristics, mission performance requirements, pilot capability and available work load capacity, and avionic automation. The primary flight control allocation problem is a major element. Some quantitative "work load" value assignable to the flight control performance requirements of the vehicle is needed. Work load is defined as the percentage of the operators time which is required for the assigned tasks during a given period. This assumes meeting mission requirements. If the work load requirements for allocated tasks is within 100 percent of the nominal individual's capability, the operator will have time to accomplish the function allocated to him. Uncommitted pilot work load capacity is available for other mandatory tasks, mission contingencies or undefined requirements, and represents the key resource available for task allocation.

Figure 1 is an illustration of the relationships involved. For a given heading hold requirement, three possible modes of flight control are available; manual unaugmented, a simple damper and a closed loop "heading hold." The heading hold requirement and associated flight conditions for the mission segment is specified in the mission. From Figure 1, which is a hypothetical data construct, we would find that manual control is not feasible, but that either damper or closed loop modes are possible. The choice of modes, and attendant equipment complexity, can then be made based upon the total work load which the flight crew has been allocated. A design decision based upon reasonable system performance and pilot work load data is then possible.

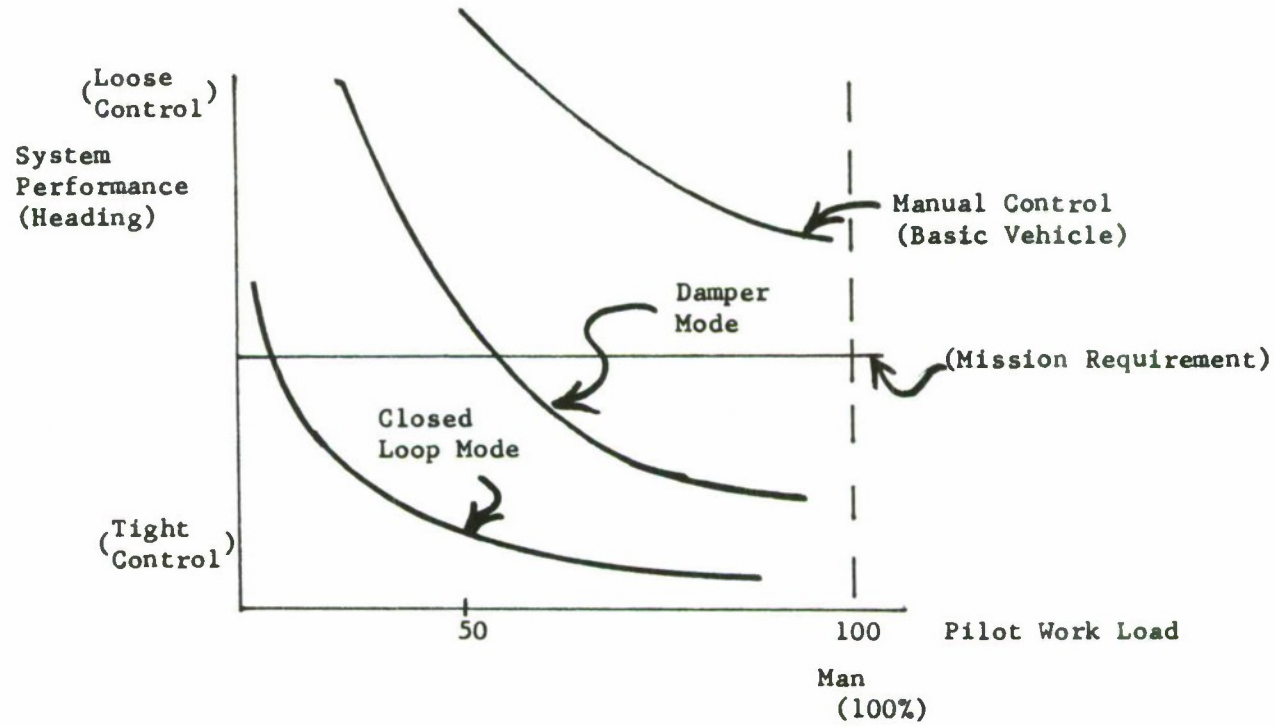


Figure 1. Functional Relationship of System Performance to Pilot Work Load and Possible Augmentation

ILLUSTRATION FROM PROJECT MERCURY

The value of work load data for flight control and the total activities related to pilot work load can be illustrated by the Mercury Space experience. The mandatory operational, flight control and housekeeping astronaut work load requirements for the Mercury configuration is shown in Figure 2. As originally planned, this was an automatic system to provide for astronaut safety and to provide for a maximum research free time in space. A second analysis studied the effects of various malfunctions. A complete manual flight control work load had been estimated. If a malfunction of the automatic flight control system occurred, a work load estimate of the same Mercury mission with total malfunction of the automatic flight control system is shown in Figure 3. Note that the in orbit work load increased to about 80 percent. This was primarily based upon an analysis indicating a 53 percent work load increase related to spacecraft attitude control in the simplest on-off mode.[1] This work was completed prior to the existence of any formal simulators or hardware.

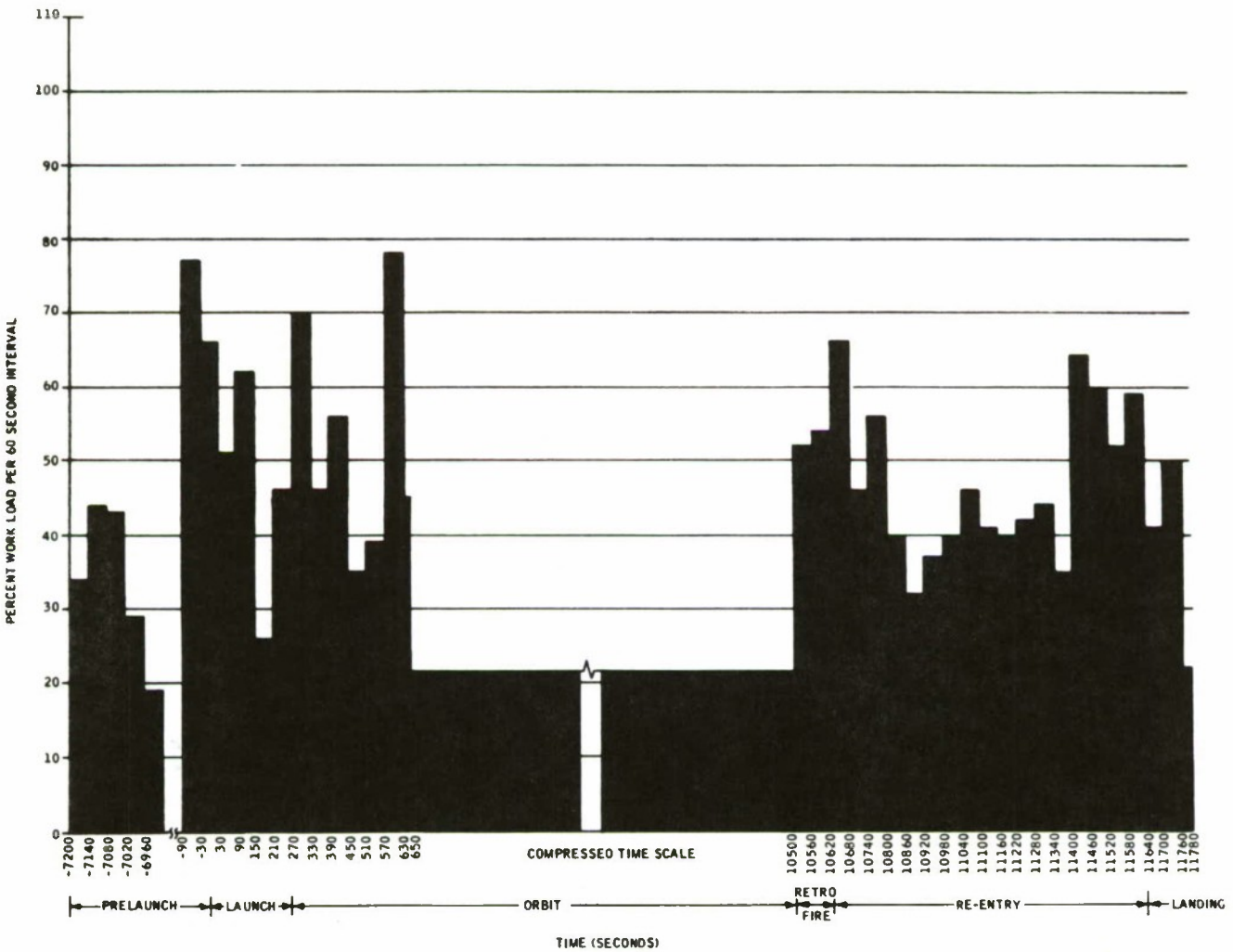


Figure 2. Normal Workload with Automatic Flight Control - Average over Sixty-Second Intervals

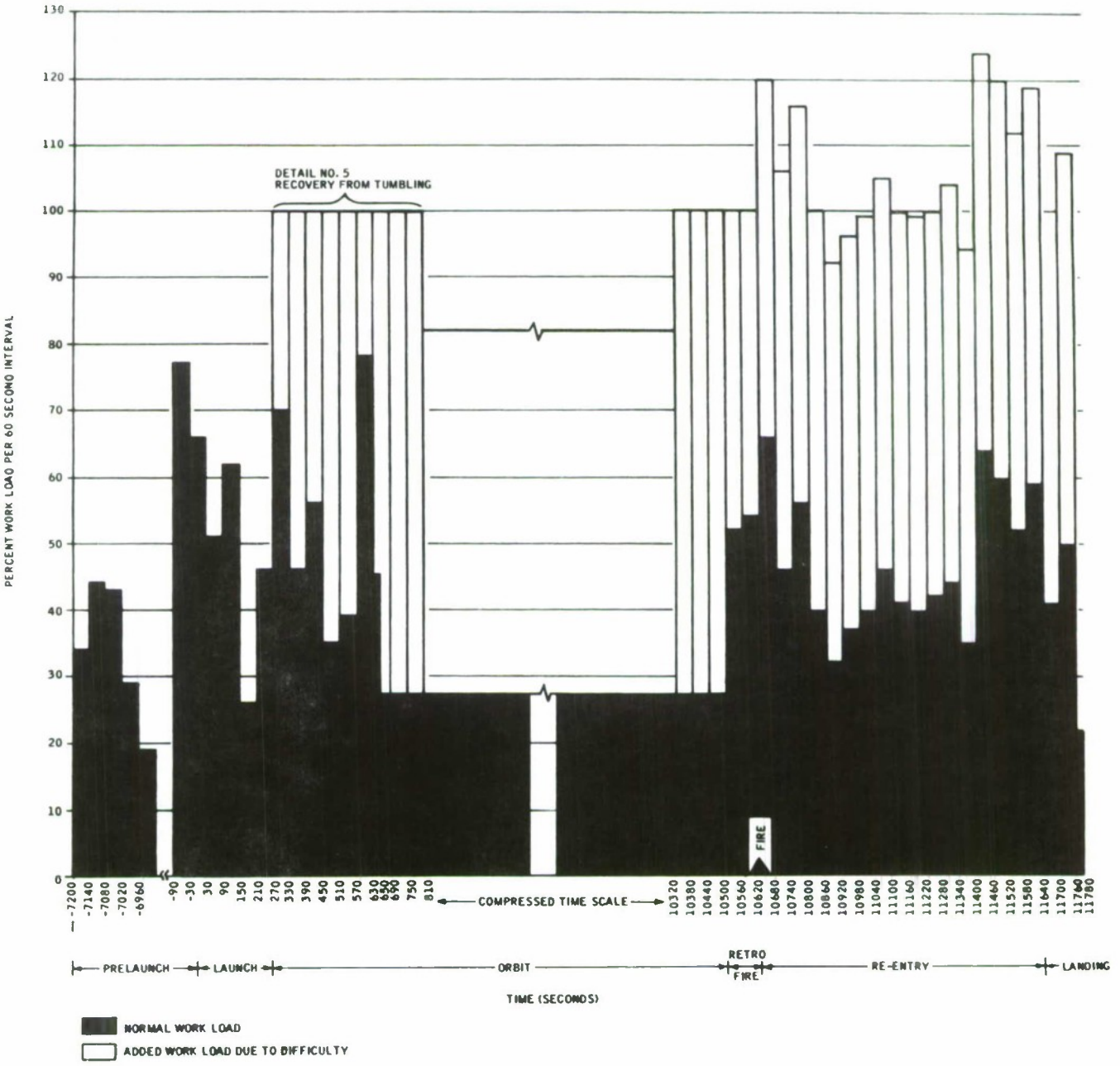


Figure 3. Full Manual Operation After Recovery from Separation

This analysis was later tested during the initial Manned Orbital Space Flight of John H. Glenn, Jr. on 20 February 1962. An operational problem related to system operation resulted in a temporary need to fly the vehicle on manual control. Some of the activity which had been scheduled was then omitted, as the early 1959 work load analysis predicted. "Because of the problems with the automatic control system..., the astronaut was forced to deviate from the flight plan to some extent, but he was able to continue all the necessary control system tasks and checks..."[2]

HANDLING QUALITIES SPECIFICATIONS

Some flight control design decisions are based upon operational experience and application of "handling qualities" guides. "Handling qualities" specifications define systems which are flyable, (within a 100 percent work load) but provide no quantitative estimate of the occupation of the pilot with his control function. There is no way of directly relating "handling qualities" to available pilot time or work load.

The interface of the pilot with his flight control problem is presently controlled primarily in terms of a "handling qualities" specification.[3] These specifications are typically quantified in terms of limit control forces, minimum and maximum breakout forces, oscillations for a single disturbance in smooth air, and simple vehicle responses for given control movements. Generally the specifications are quantitative in describing vehicle characteristics under various flight conditions, given defined external or mechanical control inputs. These specifications are however, related to the pilot by qualitative judgements which have been ellicited during flight test or by previous test data.

Handling qualitites specifications combined with simulator, iron bird, and flight test pilot evaluations can result in a "good" bird to fly or at least permit expression of handling qualities criticism in terms beyond hand waving and colorful language. The limitations of present practices are obvious in the language of the specifications and the inherent limitations of opinion polls. As examples of specification limitations, consider the following "criteria" from MIL-H-8501A (Ref. 1). (Underlines are added)

- (3.2.2) "The helicopter shall be <u>reasonably steady</u> while hovering in still air, requiring a <u>minimum movement</u> of the cyclic controls to keep the machine over a given spot on the ground...
- (3.2.5), it shall be possible <u>readily and safely</u> to bring the machine to a <u>quick</u> stop and hover.
- (3.3.9.2) During pedal fixed rolling maneuvers, there shall be <u>no objectionable</u> adverse yaw.
- (3.6.1) It shall be possible, <u>without demanding undue pilot effort</u>, to fly on instruments at all speeds, from hover to design cruise speed;

Amongst other limitations, the basic problem of the pilot -- vehicle interface is only indirectly addressed by the present "handling qualities" considerations. The nature of the problem of allocating tasks to man and machine for an effective system is illustrated by the paradox of two quotations from the December 1969 SPACE/AERONAUTICS article on the F-15.

> "One look at the combat load on the solitary F-15 pilot and it goes without saying that the total avionics design will have to be superbly automated."
>
> "Might the tradeoff on the avionics, for instance, be going against the pilot who wants more emphasis on performance to get into position to fire, less on electronics and weapon sophistication?"

The vehicle-pilot-mission interface is most certainly a key issue in the development of the display, control and sensor requirements for the pilot and co-pilot stations.

Without further information than that currently available, a design recommendation will need to provide an adequate work load performance margin by having a relatively complete flight control augmentation system or by adding a man to the crew.

APPROACHES TO FUNCTION ALLOCATION

The primary allocation of function to the man is similar in principal to the initial sizing of digital computer requirements for a new system. If all required computations, memory requirements, input-output relations, and system accuracies are considered, will the computer be capable of handling its data load and still have some reserve capacity for the undefined programming and data handling which can be anticipated? In a vehicle-pilot design, if all the defined mandatory tasks which are allocated to man in his system are accomplished, will the man be able to make his unique contributions to the mission of pattern recognition, decision making, judgement and motivation to accomplish the task? The answer to these questions for a computer require a specific "capacity" available for the mission. In the pilots case, it means he has "time" available for his functions.

This reasoning is the specific impetuous for the Honeywell introduction of "pilot work load" in the mid 1950's to the cockpit design problem.[4] Where man-machine problems have been approached from this standpoint, valid system results have been obtained. The Honeywell CF105, Mercury, and Apollo data are cited as proof of the argument. Stability or control degradation occur when flight control tasks exceed 100 percent. Tasks are omitted when work load approaches 100 percent, even though all tasks can be accomplished. This can result in poor system performance despite acceptable "handling qualities." Another conclusion verified by experience is that flight control work load does not drop to zero in manned vehicles with completely automated outer

loop control. The pilot adjusts his monitoring behavior but retains control intervention capability even when "complete automation" has been introduced.

The need remains to express vehicle flight control behavior including performance in terms of pilot work load requirements, i.e., how much of his time is required for each operation. The quantitative link needed is a relating of pilot work load to vehicle transient stability and vehicle performance. This relationship should include evaluation of the net effect of possible alternate outer-loop control modes.

Figure 5 is a block diagram of the general problem. The pilot is related to the flight control in that he executes the flight plan through mechanical inputs which he controls via monitoring of display and environmental information. The total system performance is modulated by the type of control system, the instrument display subsystem, and the type of functions which are allocated to the pilot. Handling quality specifications are directed toward providing a "reasonable" transient behavior in the mechanical system the pilot is operating.

If we simulate the complete control loop involved, including the displays and controls, we could directly evaluate the work load on the crew. The principal evaluation would involve determining whether the uncommitted work load is sufficient for the total mission tasks, or in measuring the uncommitted work load capability by some artificial measurement. This analysis method should be conducted in a full 6 degree of freedom simulator with realistic motion cues and a valid exterior view.

This type of evaluation must occur later in the system development since it requires a well defined system formulation. A reasonable engineering approximation of the total flight control problem can be obtained using simplified techniques. Techniques for early man-in-the-loop measurements have been developed which obtain a rapid assessmnet by means of a parallel man-in-the-loop and math model assessment system. Both simulations are operating on the same system simulation with a switching capability between pilot and model control modes per Figure 6.

The man-in-the-loop experiments are conducted to calibrate general performance, check the validity of the results and assist in setting the behavior of the adjustable math model. The adjustable math model is variable in gain, lead-lag characteristics, and in the display scan ratios. Since it is possible to switch from one to another, quick calibration and adjustment is possible. Once the math model is adjusted, it can be used to quickly develop data which would be time consuming if experimentally developed with a man-in-the-loop.

A comparison of data obtained by this technique in entry flight control of the Mercury Spacecraft is shown in Figure 7. This work was principally interested in the deflection point, the minimum allowable work load to control the vehicle.

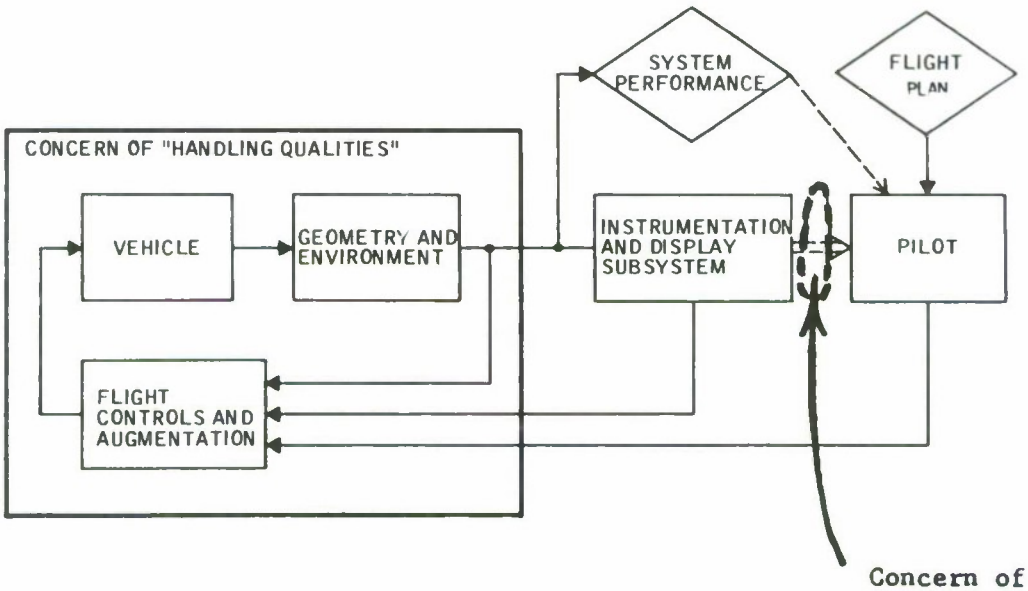


Figure 5. Flight Control Problem

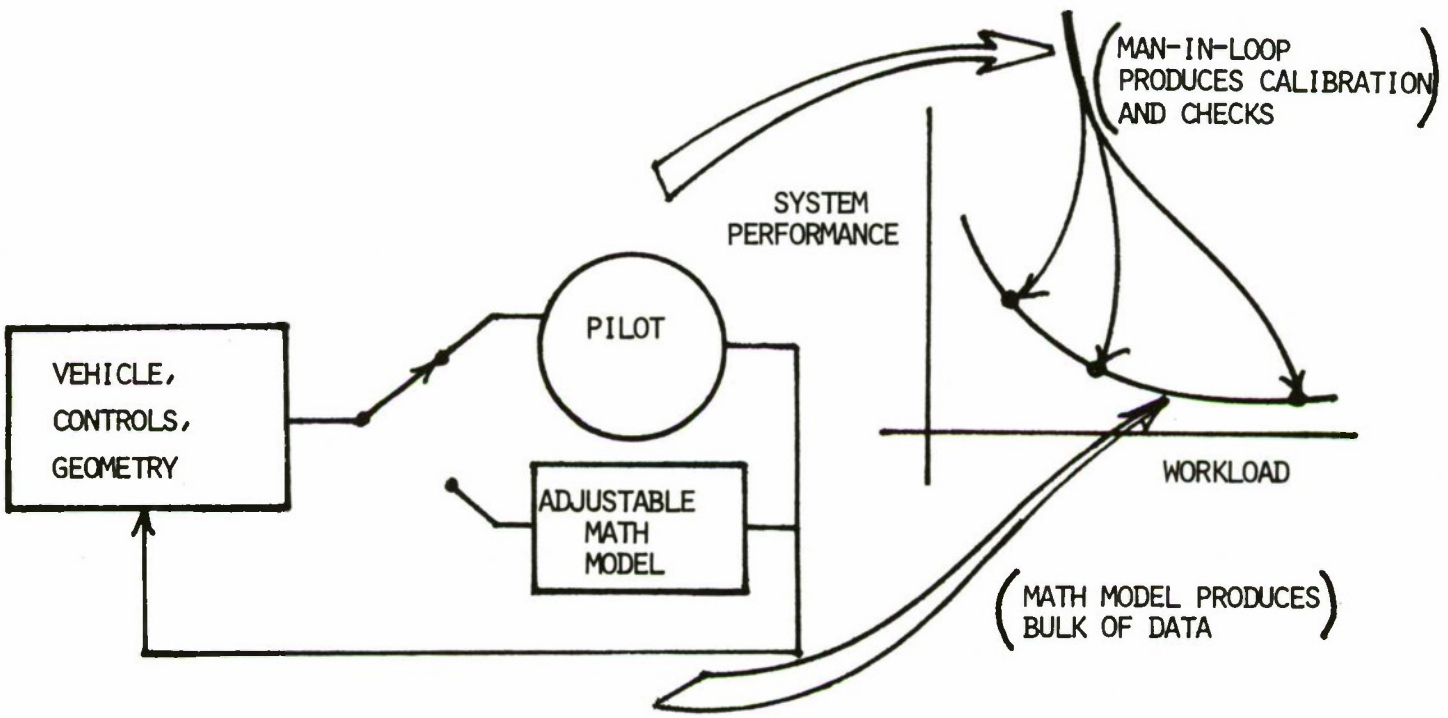


Figure 6. Block Diagram of Approach

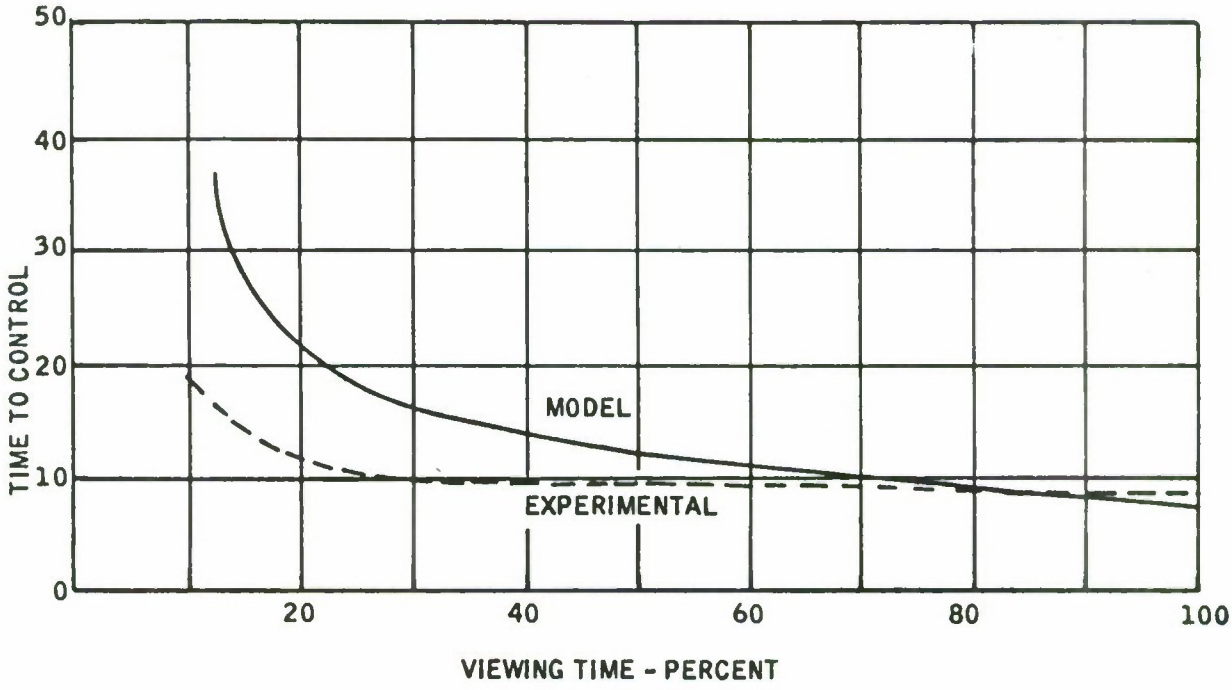


Figure 7. Comparison of Mathematical Model and Experimental Methods of Work Load Analysis

An adjustable mathematical approximation of the pilot transfer function used has been found adequate in several Honeywell studies and has yielded valuable design data. The model suits the vehicle and flight control systems descriptions just described and is adjustable to allow matching of performance. The model has been found to yield a performance estimate like highly trained and motivated subjects. The use of this invariant model allows a rapid survey of many flight conditions in a minimum amount of time. The use of a parallel man-in-the-loop simulation overcomes the limitations usually attributed to "human transfer functions" or at least defines the applicability and variance from actual operation. The duty cycle of the man is simulated by intermittantly introducing the man into the loop. The interval of loop "on" time is determined by the information content of the display subsystem, the "off" time is determined by the characteristic of the system being studied and the degree of automation employed.

Another approach to function analysis can be suggested which is based upon digital computer modeling. Consider that the pilot can be represented as a digital control system; an input-output capability, a central processing unit, and a memory. Given a general computer simulator as now available to system designers, we can introduce human performance limits into the elements of the digital computer and the associated control hardware. The output mechanical servo (human element) can be simulated with a transfer function. If we assume that human memory capability is much greater than that needed for the flight control requirement, we are left with the tracking computation time requirements and input channel capacity as variables. If we restrict human input information rates to those appropriate to conventional displays we can sample control input data to provide the flight control function. Computation of requirements can be calculated based upon auto-pilot performance iteration requirements and by assuming a nominal human computation speed.

On this basis the control system can then be operated in a typical flight mode for a given performance level. A computer program outline is shown in Figure 8. This general program could then operate in 3 degrees of a attitude control until system performance requirements are met. When performance requirements are met, the program will record time available for other tasks, the primary element of interest in task allocation. Since this system of operation has not been used before, a feasibility study is required to establish its validity.

CONCLUSION

Manned systems should be designed to provide the simplest equipment which man can use to meet mission performance requirements. The system should be designed to provide reserve crew "time capacity" to permit the man to meet the unprogramable contingencies of operation.

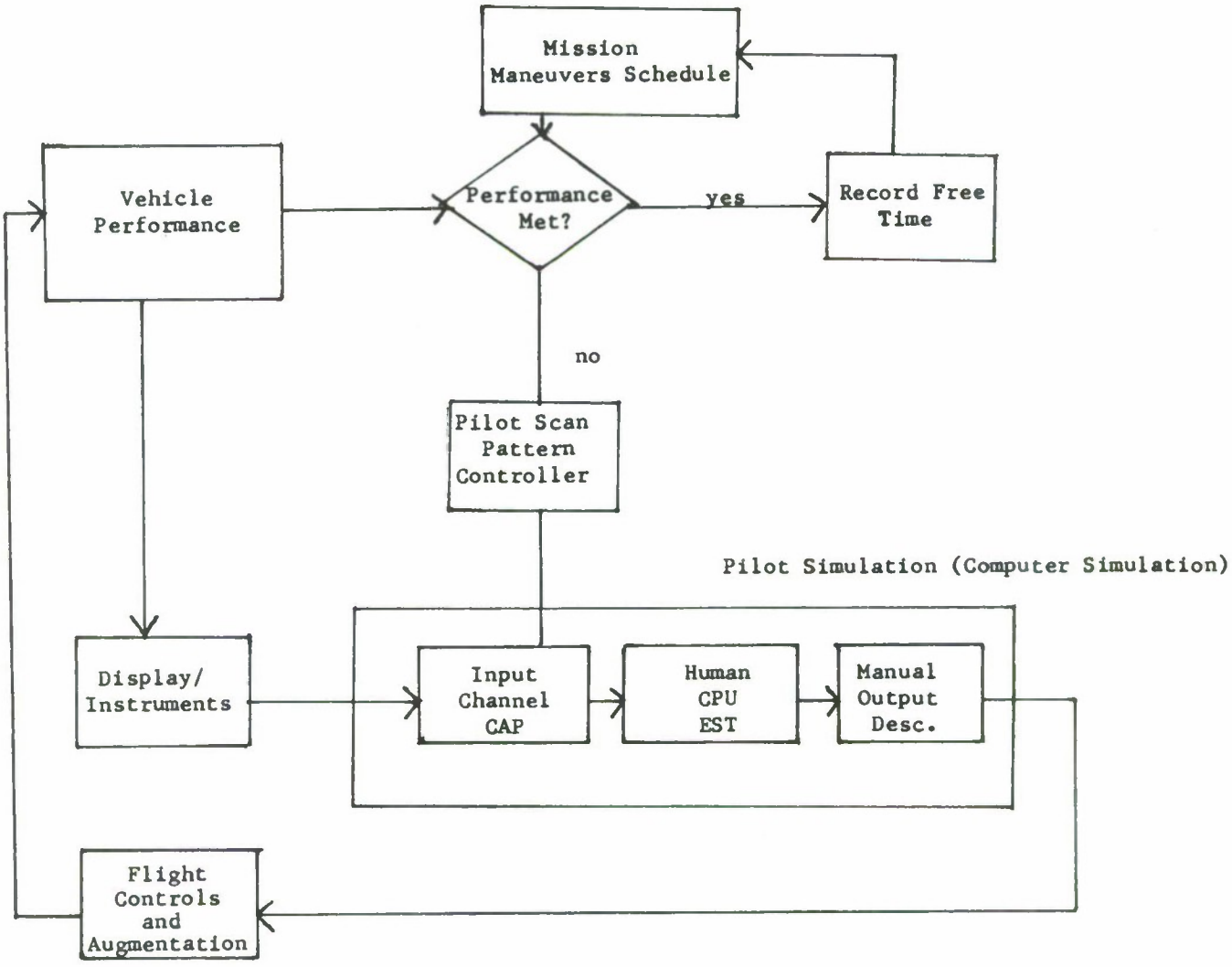


Figure 8. Digital Computer Sizing Approach to Work Load -- System Performance Measurement

BIBLIOGRAPHY

1. MH Aero Report 2315-TR2 27 August 1959

2. Results of the first United States Manned Orbital Space Flight 20 February 1962 Page 74.

3. MIL-H-8501A MIL SPEC: "Helicopter Flying and Ground Handling Qualities; General Requirements for"

4. MH Aero Report R-ED6094 October 1958

ON THE INFLUENCE OF DRUGS ON THE BEHAVIOR OF A BICYCLE RIDER.*

A. van Lunteren H.G. Stassen

Man-Machine Systems Group of the Laboratory for Measurement and Control, Department of Mechanical Engineering, Delft University of Technology, The Netherlands.
in cooperation with
M.S.H.Schlemper
Department of Psychiatry, Leyden State University, The Netherlands.

0. Abstract

The influence of some drugs on the behavior of a rider stabilizing a bicycle simulator has been studied by means of tracing the influence of the drug on the parameters of the describing functions between the input of the rider, viz. the frame angle, and the outputs, viz. the rotations of handle bar and upper body.

A pilot study with the following drugs has been executed: secobarbitali natricum; aethyl alcohol (vodka); chlordiazepoxydi hydrochloridum and perphenazinum.

It is shown that secobarbitali natricum and aethyl alcohol have a marked effect; the influences of chlordiazepoxydi hydrochloridum and perphenazinum are doubtful. In particular, the administering of vodka and secobarbitali natricum results in an increase in the time delays between input and outputs of the rider. Furthermore, it strongly acts on the remnants.

From the results obtained it can be made plausible that the upper body control is governed by hierarchically lower centers of the CNS than those which are involved in the control of the handle bar action.

1. Introduction

As a joint enterprise of the "Jelgersma Kliniek", which is the psychiatric clinic of the Leyden University, and the Man-Machine Systems Group, a study of the influence of drugs on the behavior of the bicycle rider has been started. From the view point of control engineering, the goal was to find a relation between the results obtained from the analysis of the behavior of the human operator by means of modeling techniques [1; 2; 3] and some of the psychophysiological functions involved in this control behavior.

*To be presented at the 6th Annual Conference on Manual Control, april 7-9, 1970, Wright-Patterson AFB, Ohio.

The study reported here has been executed on a bicycle simulator; with this simulator the process of stabilization has been simulated. In building the simulator, some simplifications were assumed, a detailed description of the simulator has been reported in [1; 3].

The origins of the bicycle stabilization lie in the field of those man-machine systems where the operator comes from a relatively non-selected population, and where the rider is forced to pay close attention in executing his task (for the bicycle is an unstable system). Moreover, a very attractive point is, that the rider-bicycle system is a multiloop control system as well as a multi-modality system with a time-varying character. In investigating only the stabilization phenomenon, the rider can be described by a model with one input (the angle between vertical and frame) and two outputs (the control actions of upper body and handle bar). It has been shown that the rider can be described by a model consisting of a three-term controller with a time delay [1;2;3]. The determination of the parameters in this model has been executed by means of an on-line biased parameter estimation technique as described in [3]. The block diagram of Fig. 1 shows the interaction between rider and simulator from the view point of control engineering. However, it should be

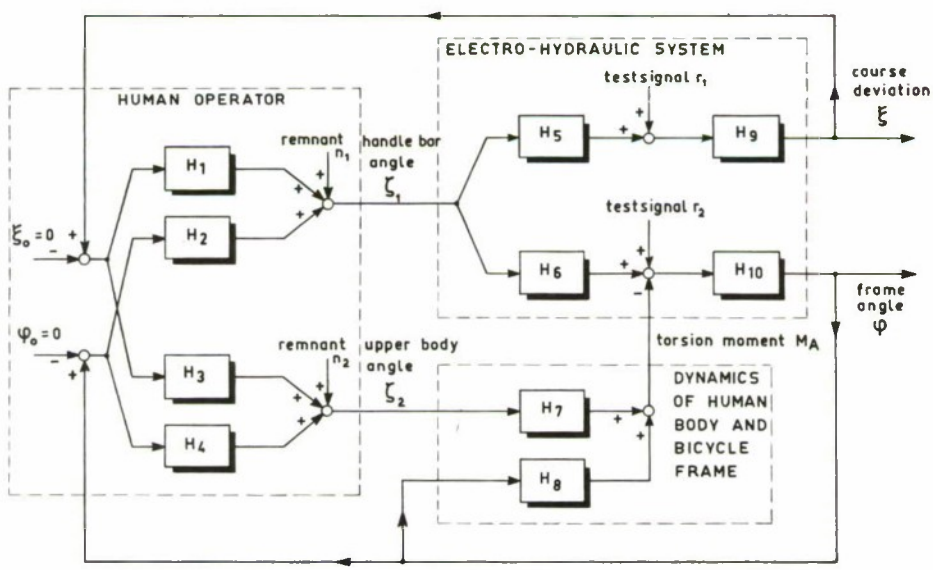


Figure 1: Block diagram of the man-bicycle system for a course following task and for a stabilization task.

very desirable if a physiological interpretation of this model could be given. Therefore the block diagram of Fig. 2 is suggested. In fact, Fig. 2 accounts more precisely

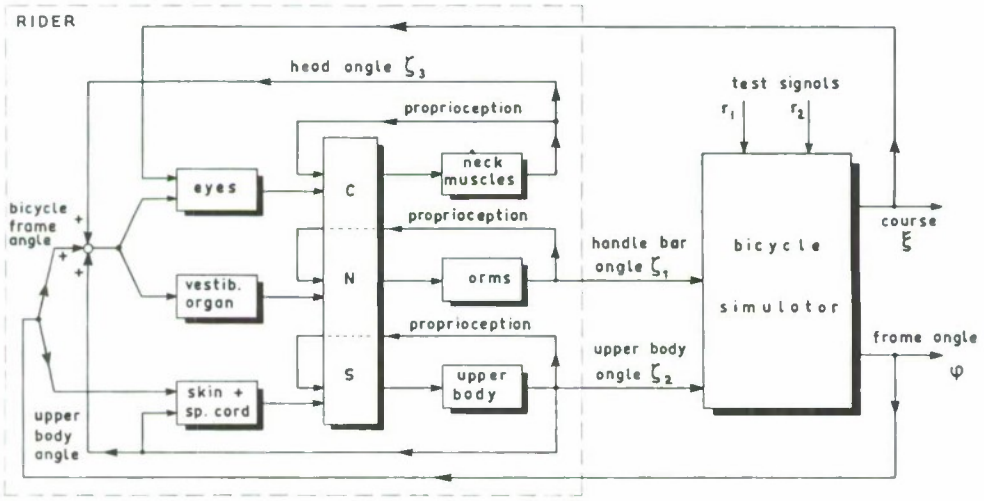


Figure 2: Block diagram of the bicycle rider based on physiological view point. Note that in this diagram also a course following task is included, and that the test signal r_1 and r_2 are simulating sidewinds and course deviations respectively.

for the stimulation of the different sensors involved. To this aim, the position of head and trunk are considered separately. The position and movements of the head in space stimulate eye and vestibular organ; the bending of neck and trunk stimulates proprioceptors in the spinal column [4]; trunk and limbs finally undergo tactile stimulation. The difficulty in handling this model will be the variety of cross couplings in the CNS. To attack this problem two important phases can be discerned. The first phase deals with the characteristics of the basic elements - sensors and pathways - which determine the behavior of the rider; the second phase is concerned with the interaction between those basic elements. A study of the basic elements can be undertaken in simple manual control tasks, so that the number of senses involved is highly reduced, for instance by ruling out the vestibular organ or the eyes. A study of the interaction between the different basic elements can be executed as published by Meiry and Young [5]. They describe experiments, where one variable was offered to the human operator using more than one sensory channel. Another

possibility is to isolate typical pathways in the CNS of the human operator, which probably can be accomplished by administering a drug having well-known effects. This means that the drugs selected must affect more or less specific centers of the CNS. In particular, those drugs which influence mainly one center should be very helpful in locating the most important pathways.

The experiments with the drugs listed in Table 1 will be reported here. A detailed description can be found in the Annual Report of the Man-Machine Systems Group [6].

specific name	trade name
secobarbitali natricum	Seconal Sodium
chlordiazepoxydi hydrochloridum	Librium
perphenazinum	Trilafon
aethyl alcohol	Vodka

Table 1: Listing of the chosen drugs.

This report also includes experiments with the neuroleptic levopromazinum; the antidepressivum amitriptylini hydrochloridum, and the stimulant amfetamini sulphas.

Finally, it should be mentioned that previous experiments have given an indication that it may be expected that the handle bar action is dependent on the system cortex and reticular formation, while the upper body motions are affected by vestibular system and cerebellum [1].

2. The drugs selected

In this section a survey of the properties of the drugs mentioned in Table 1 will be given. Here it should be noted that the effects of the drugs on the various cerebral functions need not be the same for each subject. The following review of the influence of the different drugs shows only the general trends. Furthermore, it is important to realize that certain functions in the diencephalon are related to specific transmittersubstances located at the synapses, and that particular drugs can act on a particular transmittersubstance. As a consequence hereof, drugs may act selectively on different centers in the CNS.

Secobarbitali natricum is a sleeping drug which roughly speaking acts quicker than other barbiturates. Primarily, it has a damping effect on the system cortex-reticular formation. According to van Praag [7] it mainly influences the cortex. Furthermore, in a less extent, it also acts on (a): The emotional brain. This results in a decrease of emotional sensibility, fear and tension; (b) The vestibular system and the cerebellum; if higher doses are used, it may cause ataxia and a staggering walk; (c) The gamma system. Here it leads to a lowering of the muscle tone.

Chlordiazepoxydi hydrochloridum is a minor tranquillizer acting on the same centers as secobarbitali natricum, be it with a stronger effect on the reticular formation. The influence on the neocortex is smaller. However, the emotional damping and the decrease in muscle tone are more pronounced. Higher doses may also lead to ataxia and to a staggering walk, but probably in a less degree then secobarbitali natricum.

Perphenazinum is a neurolepticum, having a damping effect on the emotional brain. Clinically, it is an effective drug to counteract hallucinations, paranoid reaction and illusionary thinking of the schizophrenic. The system reticular formation-neocortex is hardly affected: It acts on the basal ganglia leading to the Parkinsonoid syndrome (akinesia, rigidity and dyskinesia).

Aethyl alcohol has an influence comparable to that of secobarbitali natricum. It differs from most narcotics in causing a state of excitation of the subject for a longer time before drowsiness and sleep occur. Inhibiting functions are suppressed. Critical abilities and social fear decrease for instance. As with other drugs, expectations and circumstances play an important role. It should be noted that there is a difference between the effects of alcohol in an experimental condition and at a party.

3. Description of the experiments

As a first trial, the parameters of the model of the behavior of the rider executing a stabilization task were determined in three preliminary experiments with secobarbitali natricum conducted with a well-trained subject. According to Ragazzini [8] the human operator was described by a linear model consisting of a three-term controller with a delay time and, in the context of the describing function method [9] an additive remnant. If the angle φ is considered as the input to the subject informing him about the state of the bicycle, and if the variables ζ_i are consid-

ered as the control outputs of the subject (i = 1 for the motions of the handle bar; i = 2 for those of the upper body), then the form of the human transfer function will be:

$$H_i(s) = (\alpha_{i1} + \alpha_{i2}/s + \alpha_{i3}s) \exp \{- \tau_i s\}, \qquad (1)$$

where s is the Laplace operator. The parameters α_{i1}, α_{i2} and α_{i3} represent the proportional, integral and derivate constants respectively, and τ_i the delay times. To complete the describing function model of the bicycle rider the remnants $n_i(t)$ are introduced. During the experiments the bicycle simulator was used as a moped with a forward velocity of 15 km/h, while noise passing a fourth-order band-pass filter between 0.02 and 5 Hz, was introduced as test signal. As a preliminary experiment, two tests were performed. Prior to the first one, no drug was given, while the second test started half an hour after administering 100 mg of secobarbitali natricum. Three days later a test was performed half an hour after the administration of 200 mg secobarbitali natricum. During the experiments, the parameters of the model mentioned which describe the behavior of the rider were computed every minute over the previous five minutes. In consequence of these investigations a pilot study was conducted with one subject during the summer of 1968. With an exception for the alcohol, the investigations were carried out as double-blind experiments, i.e. neither the subject nor the investigators knew which drug was administered in a particular experiment. The experiments were conducted during the evening; the time between the experiments was at least one week. As a check, one experiment was conducted with a placebo. With intervals of 40 minutes six doses were administered according to one of the three following schedules A, B and C. In these schedules P means placebo (500 mg calcii gluconas) and M means medicament.
schedule A: P M P M P P;
schedule B: P P M P M P;
schedule C: P P P P P P.
The outline of the experiments is given in Tabel 2.

number	schedule	dose	medicament
1	A	2 x 8 mg	perphenazinum
2	B	2 x 100 mg	secobarbitali natricum
3	C		placebo
4	special		aethyl alcohol
5	B	2 x 10 mg	chlordiazepoxydi hydrochloridum

Table 2: List of experiments.

During the evening a number of bicycle tests were carried out. Once again, the simulator was used as moped with a forward velocity of 15 km/h, while also a similar test signal, simulating sidewinds, was applied. The total duration of one test was 10 minutes, which was split up into two periods of observation of five minutes each. This resulted in two sets of independent parameters. Before and after the bicycle experiment an additional test was carried out, viz. a reaction time experiment.

The reaction time experiment was executed as follows. An electrical stimulus was offered to the subject by means of two electrodes on the skin of his right forearm. Fig. 3 shows the set up. A low resistance between skin and electrodes was obtained by applying a conducting paste (a solution of acidi borici 1000, glycerini 280, HCl diluti 8, NaCl 400, Tragacanthi 100). In the same hand the subject held a

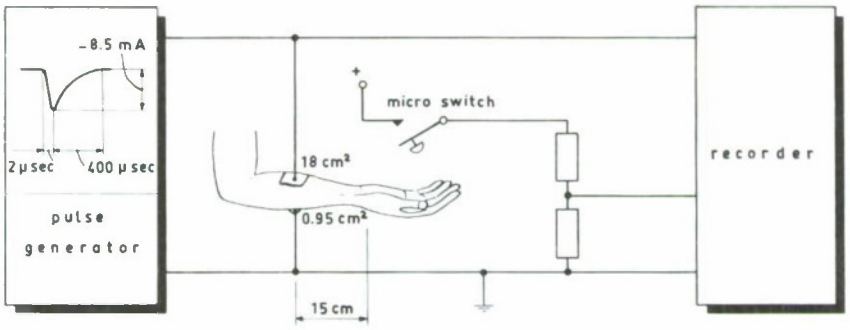


Fig. 3: Set up of the reaction time experiment.

microswitch, which had to be pressed as soon as the subject perceived the stimulus. During each test, 5 stimuli were offered at random intervals. Out of the 5 values of the reaction times measured, the highest and lowest values were omitted.

During most of the experiments the timing schedule of Table 3 was applied, see the experiments 1 up to and including 3 in Table 2. During experiment 4, the alcohol experiment, the time schedule differs only in the manner of administration. No placebos were used and the time available for taking the alcohol had to be enlarged. Every 40 minutes a blood sample was tapped. These samples were analysed by the "Gerechtelijk Laboratorium" (Laboratory of the Ministery of Justice). The experiment 5 was started five hours earlier, so that the influence of the drug used could be observed over a long term.

time	event		
13.00	warm lunch		
	during the afternoon no coffee or alcohol		
17.00	light meal of bread		
	drug or placebo	bicycle simulator test	reaction time test
19.50	1		
20.15			1A
20.20-20.30		1	
20.30	2		1B
20.55			2A
21.00-21.10		2	
21.10	3		2B
21.35			3A
21.40-21.50		3	
21.50	4		3B
22.15			4A
22.20-22.30		4	
22.30	5		4B
22.55			5A
23.00-2310		5	
23.10	6		5B
23.35			6A
23.40-23.50		6	
23.50			6B

Table 3: Schedule of the tests in one experiment.

Before starting the experiments with drugs a number of tests were made by the subject under study, so that he could be considered as well-trained.

The parameter estimation was accomplished with a 4k PDP8.

4. Results

The outcomes of the preliminary experiment with secobarbitali natricum are given in Fig. 4. The results suggest that this drug has a marked influence on the behavior of the rider. The delay times and the magnitude of the remnants seem to be affected. The Fig. 4 items d and j show that with a dose of 100 mg primarily the delay time of the handle bar action increases, while that of the upper body movement hardly changes. However, the delay time of the upper body movement is affectd only after a further increase in the dose to 200 mg, while the delay time of the handle bar action shows no further augmentation. The relative remnantsbehave just in an opposite manner. The relative remnant of the upper body action increases from 0.4 till even 0.9, while that of the handle bar action decreases from 0.3 till a value as small as 0.15.

The outcomes of the experiments listed in Table 2 are given in the Figs. 5 up to 8 inclusive. To allow a judgement on the significance of the changes observed, the standard deviations σ of all quantities under normal conditions have been determined. The standard deviations were calculated from the first test of each drug experiment as well as from three introductory experiments. In all figures the mean value and the confidence limits for 2σ are indicated.

As the coefficients of the integrative action were always close to zero, these parameters are left out in the figures. Due to a malfunctioning of a part of the interface during the experiment with secobarbitali natricum, no reliable parameter estimation could be executed. Therefore, only a few quantities could be determined; they are not given in this paper. In general, the trend was in accordance with that of the preliminary experiment.

During the experiments the subject was observed by a psychiatrist to check probable deviations from his normal behavior.

5. Conclusions and further research

The table 4 summarizes the results of the experiments. In this table is indicated in what way the quantities are affected by the drugs administered. The rows 1 up to and including 5 indicate the experiments of the pilot study. A separate row is added representing the preliminary experiment with secobarbitali natricum. The sign "0" indicates that no significant change was found. A "+" or "-" sign means respectively an increase or decrease in the related quantity beyond the 2σ-limits, while a "++" or "--"

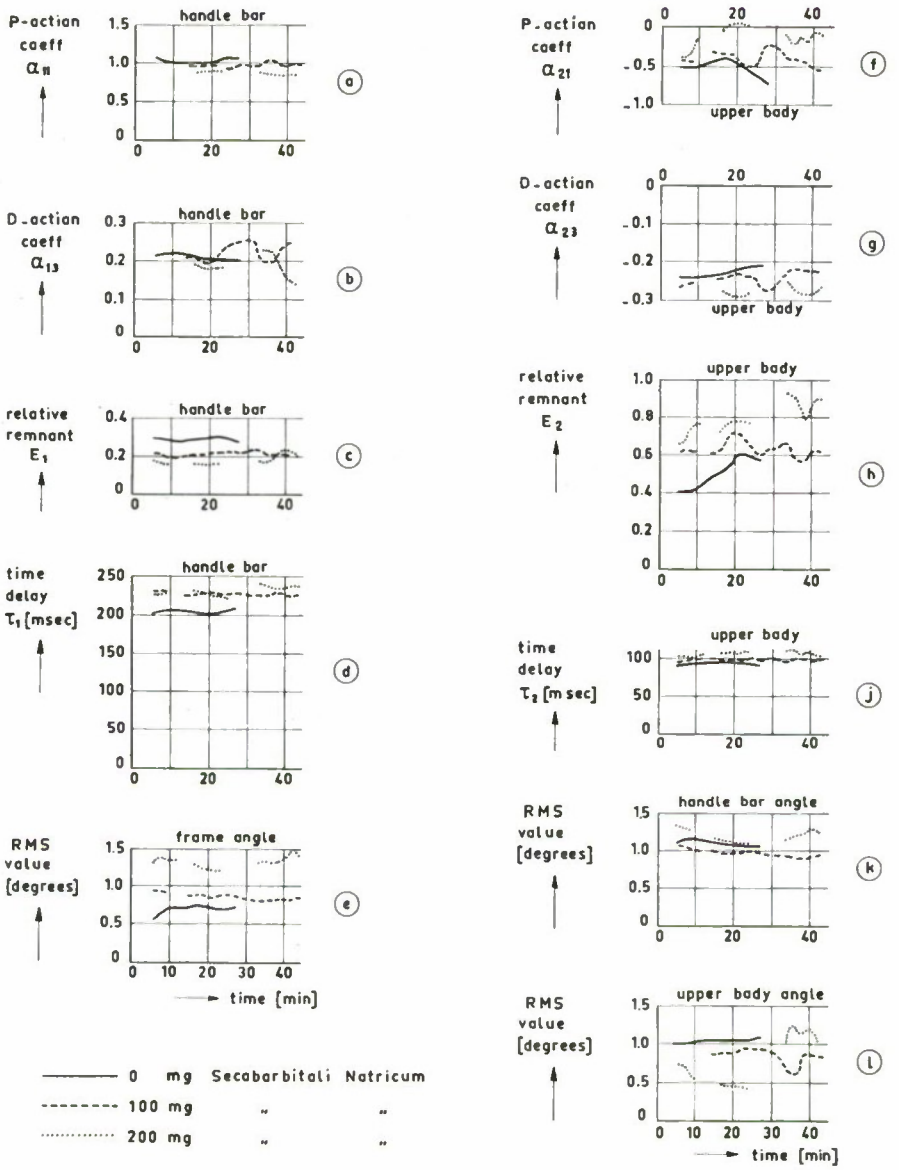


Figure 4: Results of a preliminary experiment with secobarbitali natricum.

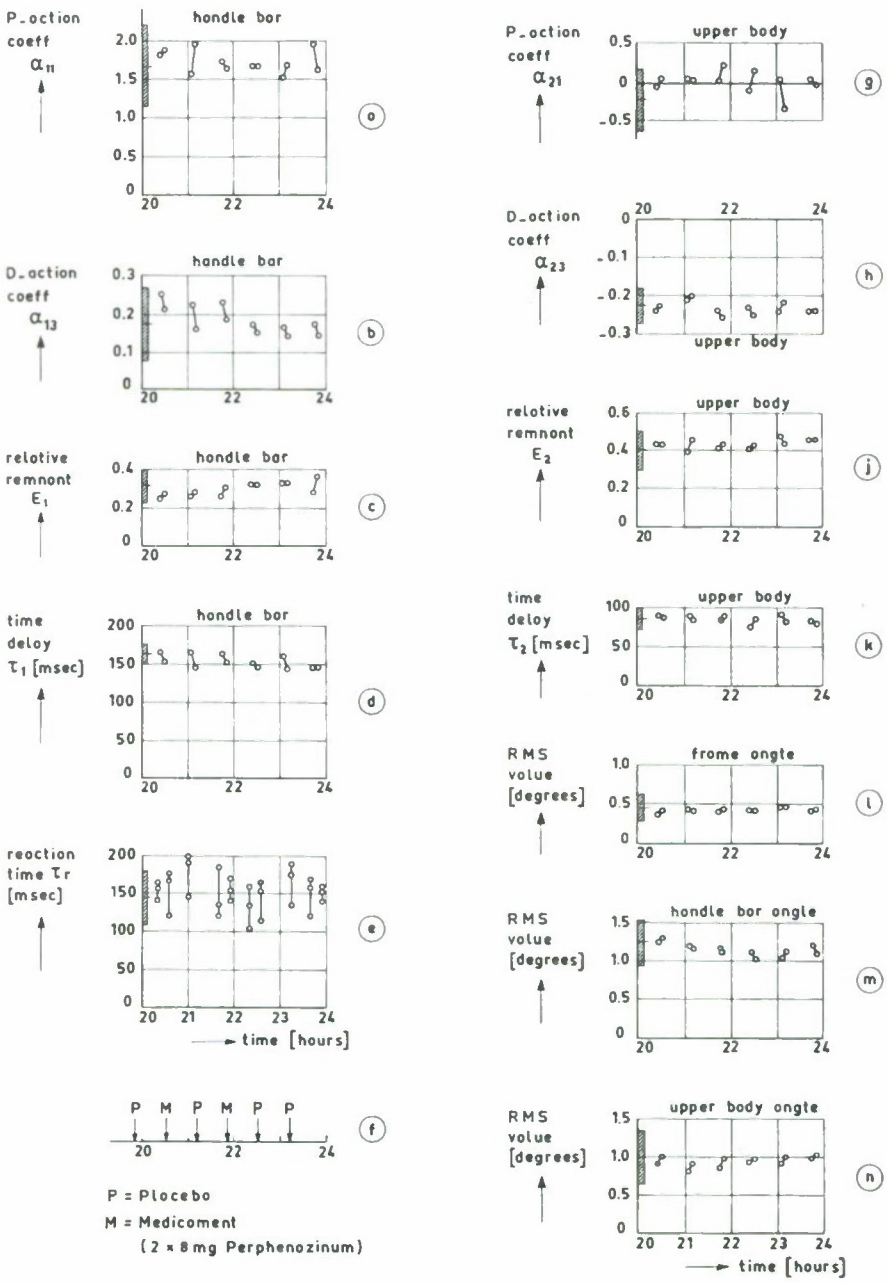


Figure 5: Results of the experiment 1 of the pilot study.

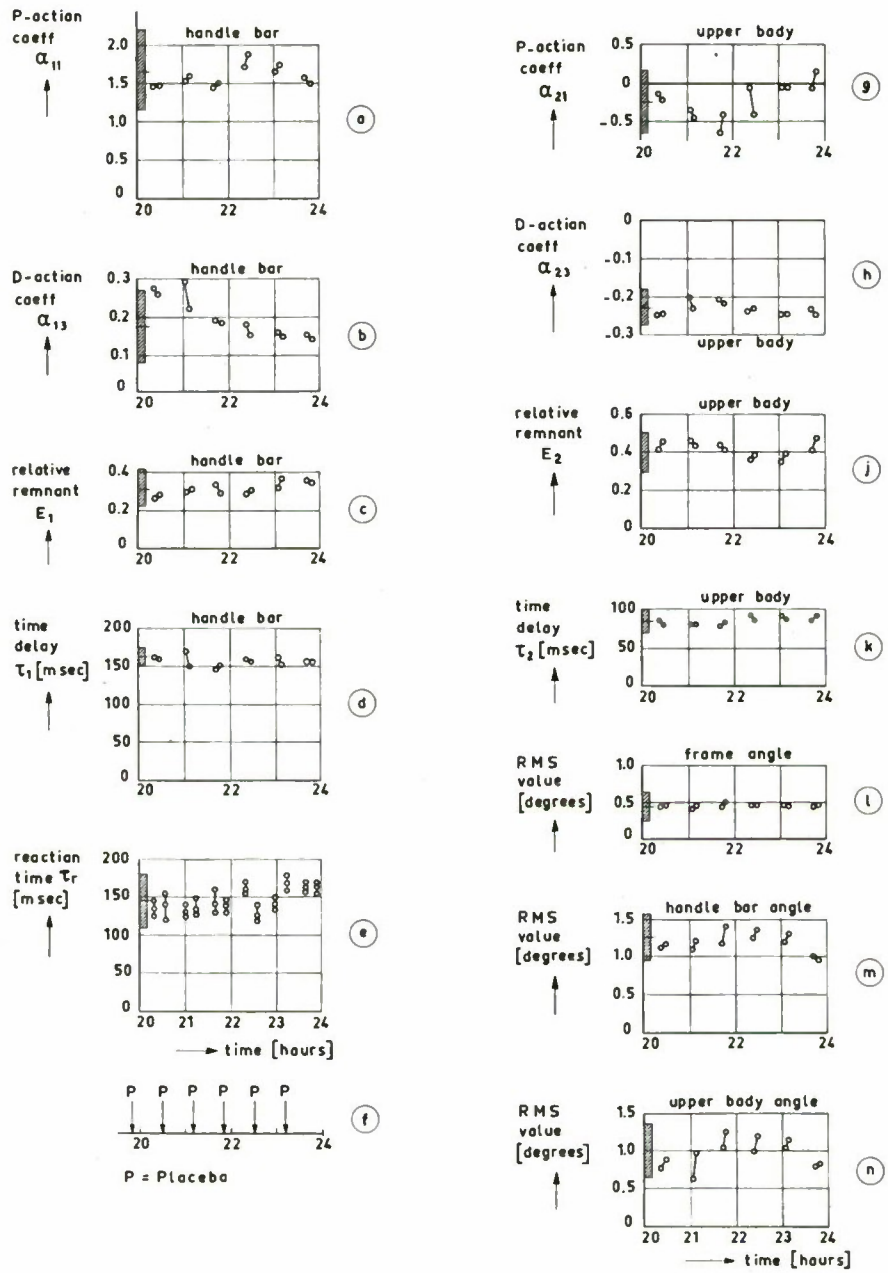


Figure 6: Results of the experiment 3 of the pilot study.

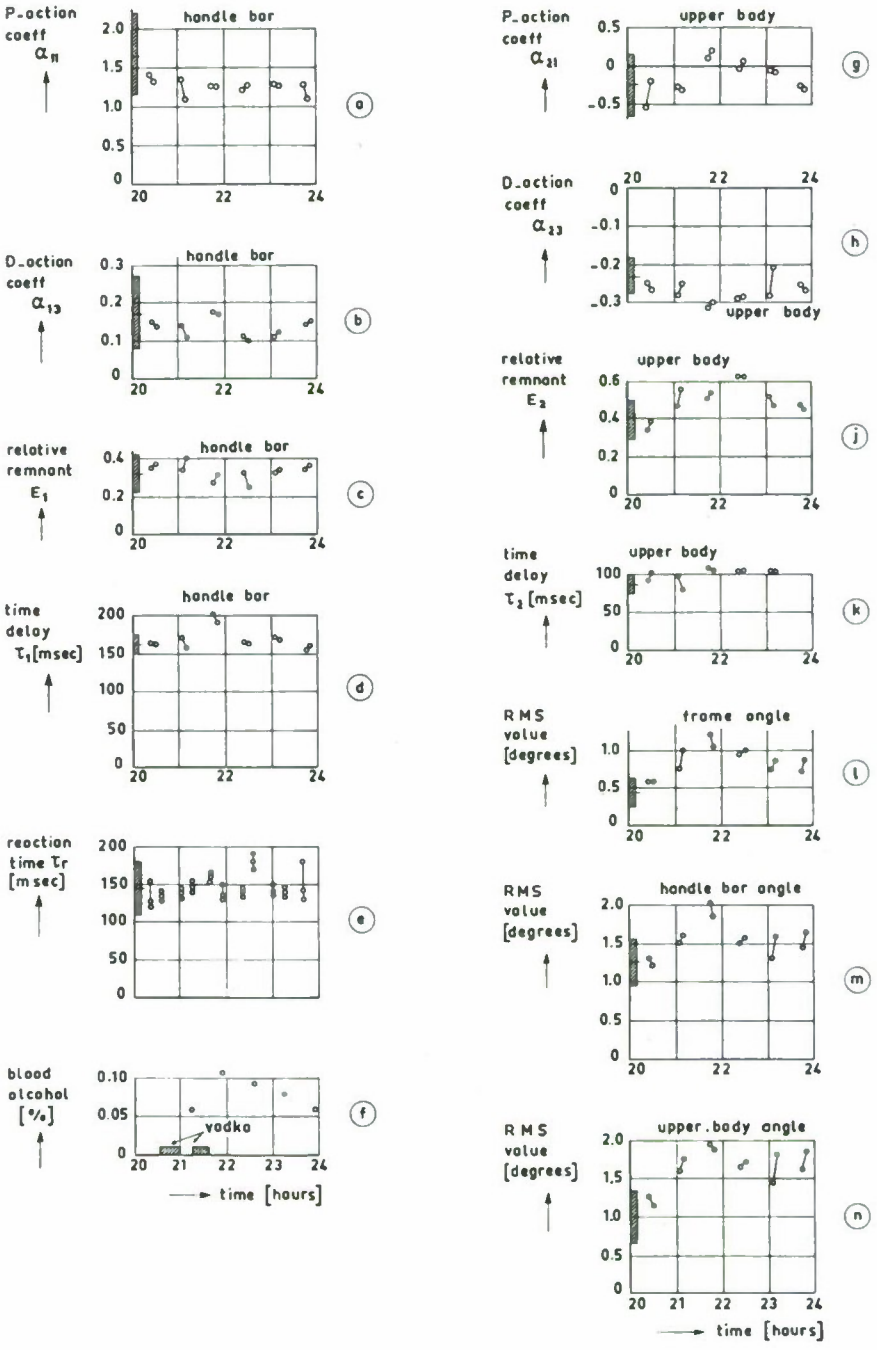


Figure 7: Results of the experiment 4 of the pilot study.

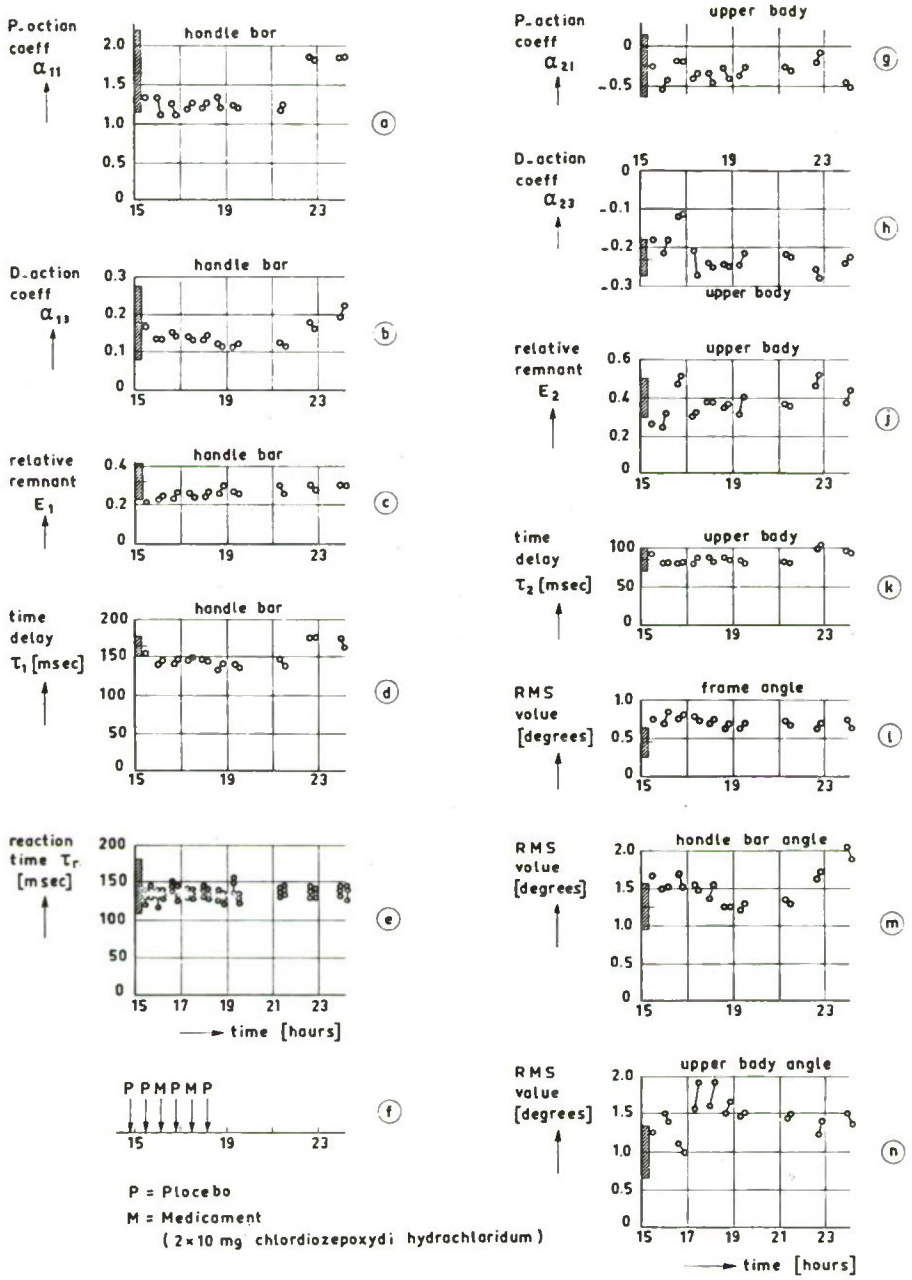


Figure 8: Results of the experiment 5 of the pilot study.

sign means an increase or decrease significant with respect to the 3σ-limits. The table 4 shows that vodka and secobarbitali natricum have a marked effect; there is an influence of chlordiazepoxydi hydrochloridum on the RMS values of the control actions, while perphenazinum and

quantities measured / drug	number	handle bar				upper body							
		α_{11}	α_{13}	E_1	τ_1	α_{21}	α_{23}	E_2	τ_2	τ_r	σ_φ	σ_{ζ_1}	α_{ζ_2}
perphenazinum	1	0	0	0	-	0	0	0	0	0	0	0	0
placebo	3	0	0	0	0	0	0	0	0	0	0	0	0
vodka	4	0	0	0	++	0	-	++	+	0	++	++	++
chlordiazepoxydi hydrochloridum	5	0	0	0	-	0	++	0	0	0	0	++	++
secobarbitali natricum	pr. exp.	0	0	-	++	0	0	++	+		++	0	-

Table 4: Influence of the different drugs on the quantities measured.

the placebo do not change the performance of the subject. Only during the alcohol experiment a blood analysis has been executed. For the other drugs such an analysis is rather complicated. For those quantities which change significantly, the relation between these quantities and the blood-alcohol content are given in Fig. 9. This figure shows the well-known effect that the influence of a certain percentage alcohol in the blood during the increasing phase of the blood-alcohol curve is stronger than during the decreasing one. This phenomenon can be considered as being a certain degree of adaptation of the body to the alcohol. It also should be noted that changes of delay times and remnants are in accordance with those observed in the preliminary experiment with secobarbitali natricum, viz. an increase in both the delay times, an increase in the remnant of the upper body action, and a decrease in the remnant of the handle bar action.

As was mentioned in the section "drugs selected", alcohol and secobarbitali natricum are known to have similar effects. Furthermore, the results in Fig. 9 show that alcohol causes

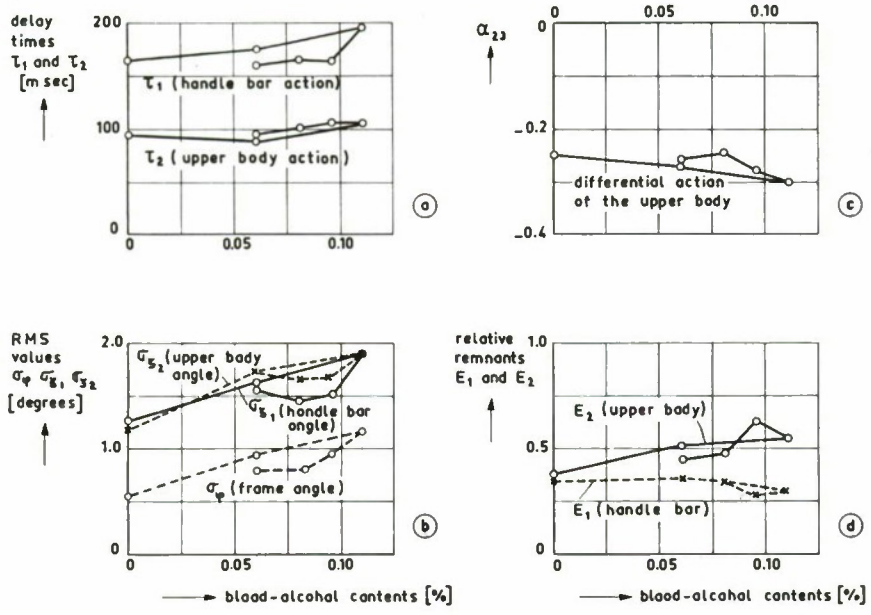


Fig. 9: Regression lines between some of the quantities measured and the blood-alcohol contents.

changes in the quantities observed which are strongly correlated. In the other experiments, where no blood-drug contents are known, these mutual relations might give some information. However, in most of the experiments such a relation can hardly be found. Only the relations between the RMS-values of, on the one hand, the upper body and, on the other, the handle bar angle and the frame angle sometimes exist. This is illustrated in Fig. 10.

The delay time of the handle bar action is always about twice that of the upper body. This suggests, that the upper body control is governed by hierarchically lower centers of the CNS than those which are involved in the control of the handle bar action. This implies that the strategy of the bicycle rider in controlling the handle bar will be more flexible in comparison with the upper body control strategy. This poorer flexibility of possible preprogrammed actions governed by the lower centers in the CNS, probably has something to do with the shorter delay time found. If the functioning of the lower CNS-centers is influenced, this may result in a malfunction of the control actions governed by these centers. However, in contrast to the lower CNS-centers, the higher ones possess a greater

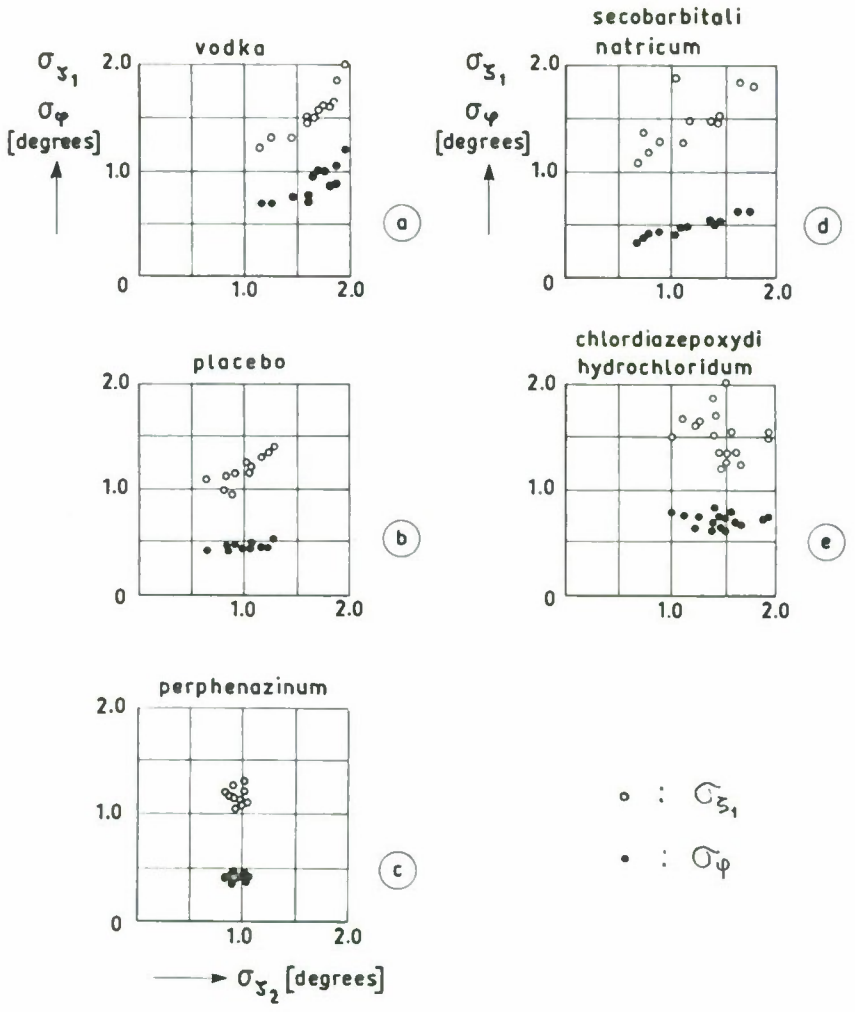


Fig. 10: Relations between the RMS values σ_{ζ_1}, σ_φ and σ_{ζ_2} for different drugs.

flexibility, so that possible disturbances of their functioning probably can be compensated up to a certain degree. If it is assumed that the upper body control is dominated by the vestibular-cerebellar system, and if it is assumed that the handle bar actions are governed on a cortical level, the influences of Vodka and secobarbitali natricum on the rider's behavior can be explained as follows:
The deterioration of the vestibular-cerebellar system results in an increase in the delay time τ_2 and in the relative remnant E_2. The increase in this remnant means that the actions of the upper body are less adequate. Although the function of the neocortex is also influenced,

which is shown by the increase in the delay time τ_1 of the handle bar action, the subject is able to compensate for this influence. This can be concluded from the decrease in the relative remnant of the handle bar action. The increase in the RMS-values of the frame angle, handle bar angle and upper body angle might be caused by an increase in the upper body angle, which results in an increase in the frame angle. In order to keep the system under control a larger value of the handle bar angle will be necessary.

Fig. 10 shows, that for most of the drugs a correlation exists between the RMS-values of frame angle, handle bar angle and upper body angle. There is a marked difference with respect to these quantities between chlordiazepoxydi hydrochloridum (Fig. 10e) and perphenazinum (Fig. 10c). Chlordiazepoxydi hydrochloridum has a very large dispersion while perphenazinum shows that the quantities mentioned remain practically constant. However, in the reaction time experiments the outcomes behave just the opposite way. Whether these outcomes are accidental or not, remains an open question.

In the further research the following items should be taken into account:

A. The task of the rider was probably too easy to be influenced by the majority of the chosen drugs. Extension to a double task, which requires more attention, will be necessary, It is intended that the bicycle rider has to follow a prescribed course as the primary task, while the balancing of the bicycle will provide a secondary work load. This load may be increased by changing the dynamics of the simulator. In the latter case, the analysis of the behavior of the rider by means of parameter estimation techniques, will be more difficult because the interaction between the two inputs and two outputs of the rider has to be taken into account, and thus the model will be more intricate. Furthermore, an unbiased estimate of the parameters should be desired in order to obtain adequate quantities, which might be used for a physiological interpretation.

B. Knowledge of the drug content in the blood might give much additional information. However, even if it is possible to obtain this knowledge from blood samples for each of the drugs used during a particular investigation, it should be realized that taking blood samples may influence the condition of the subject.

6. Acknowledgements

The authors wish to express their gratitude to Prof.dr W. Froentjes of the "Gerechtelijk Laboratorium" for his contribution in the analysis of the blood samples

of the alcohol experiment. The medical staff of the Delft University Health Department are thanked for their cooperation in giving the subjects in the bicycle experiments a preliminary medical examination.

7. References

1. Lunteren, A. van; Stassen, H.G.: Investigations on the characteristics of a human operator stabilising a bicycle model, Intern. Symp. on ergonomics in machine design, Prague (1967), p. 27.
2. Lunteren, A. van; Stassen, H.G.: On-line parameter estimation of the human transfer in a man-bicycle system, IVth IFAC congress, paper 70.3, Warsaw (1969), p. 17.
3. Lunteren, A. van; Stassen, H.G.: On the variance of the bicycle rider's behavior, 6th Annual Conference on Manual Control, Wright Patterson AF Base (1970), p. 22.
4. Cohen, L.A.: Role of eye and neck proprioceptive mechanisms in body orientation and motor coordination, Journal of Neurophysiology 24 (1961), pp. 1-11.
5. Young, L.R.; Meiry, J.L.: Manual Control of an unstable system with visual and motion cues, IEEE International Congr. Rec., vol. 13 pt. 6, (1965), pp. 123-127.
6. Lunteren, A. van; Stassen, H.G.: Annual Report of the Man-Machine Systems Group, Laboratory for Measurement and Control, Dpt. of Mech. Eng.,Delft University of Technology, Delft (1970).
7. Praag, H.M. van: Psychofarmaca, van Gorcum & Cie, Assen (1966), p. 228.
8. Ragazzini, J.R.: Engineering aspects of the human being as a servomechanism. Unpublished paper presented at the Am. Psychol. Assn. Meeting (1948).
9. Booton, R.C.: The analysis of non-linear control systems with random inputs, Symp. on non-linear circuit analysis, Polytechn. Inst. of Brooklyn (1953), pp. 369-391.

INDEPENDENT CHANGES IN VISUAL AND KINAESTHETIC FEEDBACK

Leslie Buck
Control Systems Laboratory
National Research Council of Canada.

ABSTRACT

Experiments on stimulus-response compatibility demonstrate that performance is degraded by changes in feedback relations. A technique is reported which attempts to localize this effect by varying kinaesthetic and non-kinaesthetic feedback independently.

Experimental results show incompatibility leads to increases in movement time, as well as reaction time, showing that movement execution is affected by feedback changes. This is so whether the change is essentially one of kinaesthetic feedback or of visual feedback. In the former case, however, recovery to previous performance levels is more rapid, suggesting that kinaesthesis is the more important source of feedback, even though not the only one used.

INTRODUCTION

The role of feedback in the control of movement has been recognised in 'guided' as opposed to 'ballistic' theories of movement. Guided theories hypothesize that a movement is under more or less continuous control during the whole course of execution, and not only at its initiation and completion. The question of how this control is mediated is another matter, although some source of information is obviously required, of which visual and kinaesthetic feedback are two possibilities. Identifying this source is of some importance for advocates of guided theories since the long latencies associated with visual feedback (in relation to the time scale of the movements themselves) are adduced in evidence against such theories. On the other hand such data have also been used by Gibbs (1965) as evidence that if feedback is used then it must be kinaesthetic rather than visual.

The relation normally found between different forms of feedback has led to the concept of stimulus-response compatibility. When an operator makes a control action he not only receives kinaesthetic feedback from his musculature, which tells him about the position and movement of his limbs and other response apparatus. He also gets feedback from, for example, hearing the sound of his own voice, seeing

his hands move, seeing the movement of an instrument directly coupled to the control, or seeing or sensing in some other way the change in the process brought about by the control action. As a consequence changes in one form of feedback, of a given kind or degree, are associated with changes in another form, of another kind or degree. One might even argue that stimulus-response compatibility would be better named feedback-feedback compatibility.

According to the guidance theory of movement, a response, once it has been selected, is controlled by comparing feedback arising from its execution with a learned pattern of feedback, and correcting for error between the two. A skilled movement has associated with it a learned pattern like this, and part of the learning process consists of acquiring the pattern.

Some indication of the role of feedback has been demonstrated by reversals of relations whereby a given kinaesthetic sensation is now associated with the previous visual sensation in mirror image, and vice-versa, and Bernotat (1969) has extended this technique to include phase shifts other than 180°. In this procedure the phase shift has taken place in, as it were, a free-floating situation: that is to say one cannot specify whether it was the kinaesthetic feedback that was changed, or that in the other sensory mode. This is because the two are considered relative to each other with no external reference point.

If an external reference could be incorporated we might study independent changes in kinaesthetic and non-kinaesthetic feedback and compare their effects relative to each other and a control situation of no change. Thus we would give the subject a task in which he monitors his movements by reference to kinaesthetic feedback but not by reference to non-kinaesthetic feedback - since this has been changed - and vice-versa. We have developed a procedure which, we think, allows us to do this.

EXPERIMENTAL METHOD AND DESIGN

The principle of this method is to teach the subject a pattern of movements which provides him with patterns of kinaesthetic and non-kinaesthetic (in our case, visual) feedback which he can learn. We use a subject-paced pursuit tracking task (Figure 1) (Buck et al, 1970) with a compatible relation between control (steering wheel) and display (oscilloscope), in which the target moves step-wise in an overall left-right direction (Figure 2) between five horizontally aligned positions. The task is carried out in complete darkness so that visual feedback comes only from the pursuit spot. Between runs the display is blanked and the subject releases the wheel.

During the course of nine training runs the subject learns the

patterns followed by his hands and by the pursuit spot, and in the six experimental runs he uses one or other of these to monitor his movements. Thus in condition V the subject is presented with targets which follow a mirror image (that is, right-left) pattern, but the control-display relation is inverted so that he has to make exactly the same wheel movements as in training, producing the same kinaesthetic feedback. In condition K the control-display relation is inverted but the same left-right target pattern is presented so that he has to make the same pursuit spot movements producing the same visual feedback. In condition C (control) no change is made. A fourth condition VK uses the mirror-image pattern with direct control-display relation so that both visual and kinaesthetic feedback change.

Condition	Target Pattern	Control-display Relation
C	left-right	direct
V	right-left	indirect
K	left-right	indirect
VK	right-left	direct

In our experiment ten subjects were tested under each of the four conditions, forty subjects in all. Each run consisted of five consecutive sweeps of the ten step target pattern, and at the end the subject was told his number of errors for the run. An error is defined as a movement initiated in the wrong direction away from the target.

In measuring performance we distinguished between reaction time - the length of the period during which the subject is selecting and preparing his responses - and movement time which is the time taken to execute the response. (Movement time as defined here terminates when the new target is reached and does not include the time spent correcting for overshoots.) In the experimental runs of conditions V and K subjects had to select and prepare responses knowing that the control-display relation had been reversed. We predicted therefore an increase in error rate and mean reaction time following the change. However, since during the reaction time period no feedback is received by the subject, we did not expect groups V and K to differ in the extent of the increase.

Once a response had been correctly selected and the movement correctly initiated feedback became available to control the movement. From the point of view of ballistic theories of movement however this is not important, and therefore one would not predict any change in mean movement time as a consequence of the change in the control-display relation. From the point of view of guided theories, on the other hand, this feedback is very important and some increase in mean movement time would result because of the discrepancy between previously learned, and actual, patterns of feedback.

Whether or not equal impairment is predicted for both groups V and K depends on the theoretical importance of differential changes in visual and kinaesthetic feedback. Our hypothesis emphasizes the predominant role of kinaesthetic feedback and we predicted therefore a higher increase in mean movement time under condition K than under condition V since, in the latter condition, actual kinaesthetic feedback remained (assuming responses were correctly selected) the same as the previously learned pattern.

RESULTS AND DISCUSSION

The experimental results confirm that error rates and mean reaction times for correct responses increased during the experimental runs of conditions V and K and that there was virtually no difference between the two groups (Figures 3 and 4). Furthermore, there was an increase in mean movement time under these two conditions (Figure 5). We may conclude therefore that the change in the control-display relation affected both the selection of the response and the execution of the movement.

Regarding the effects of the different changes in feedback, the results do not confirm the hypothesis: allowing for overall differences during the training runs, both groups were equally degraded at the first experimental run. There was however one point of difference between them. During the six experimental runs the mean movement times of group K fell to the level previously attained during the training runs. In group V on the other hand there was less improvement over this period and at the end mean movement time was still considerably higher than at the last of the training runs.

This recovery effect indicates that our technique had some success in producing independent changes in visual and kinaesthetic feedback. The consequences of the changes were not, however, as we predicted and some re-appraisal of our hypothesis is required.

The evidence of the initial increase in movement time supports the hypothesis that movements are under continuous control, and the increase can be ascribed to the introduction of an incompatibility between the visual and kinaesthetic feedback associated with the movement. Subsequent improvement depends upon the resolution of this incompatibility, that is to say the learning of a new relation between the two kinds of feedback. This learning proceeds more rapidly in that condition where a new kinaesthetic pattern is being associated with an old visual pattern: in the reverse case learning occurs also, but less rapidly.

The greater facility with which kinaesthetic learning proceeds may point to its greater importance in the control of movement. At the same time our results suggest that the control function cannot be ascribed to kinaesthesis alone and that some role must be allocated to visual feedback. One possibility is that the various forms of feedback available

in a given situation may serve to monitor each other: while one form (kinaesthesis perhaps) provides a first-order indication of error in execution of the movement, other forms in turn give indications of error in the first error indication. We may hypothesize that the introduction of new relations between the different forms of feedback disrupts this process and lengthens movement time.

REFERENCES

1. Bernotat, R. (1969) Rotation of visual reference systems and its influence on control quality. Proceedings of the Fifth Annual NASA-University Conference on Manual Control.

2. Buck, L. Hyde, F., Leonardo, R. (1970) A Subject-Paced Step-Input Pursuit Tracking Task. LTR-CS-28, Control Systems Laboratory, Division of Mechanical Engineering, National Research Council, Ottawa.

3. Gibbs, C.B. (1965) Probability learning in step-input tracking. British Journal of Psychology, 56, 233.

Experimental apparatus

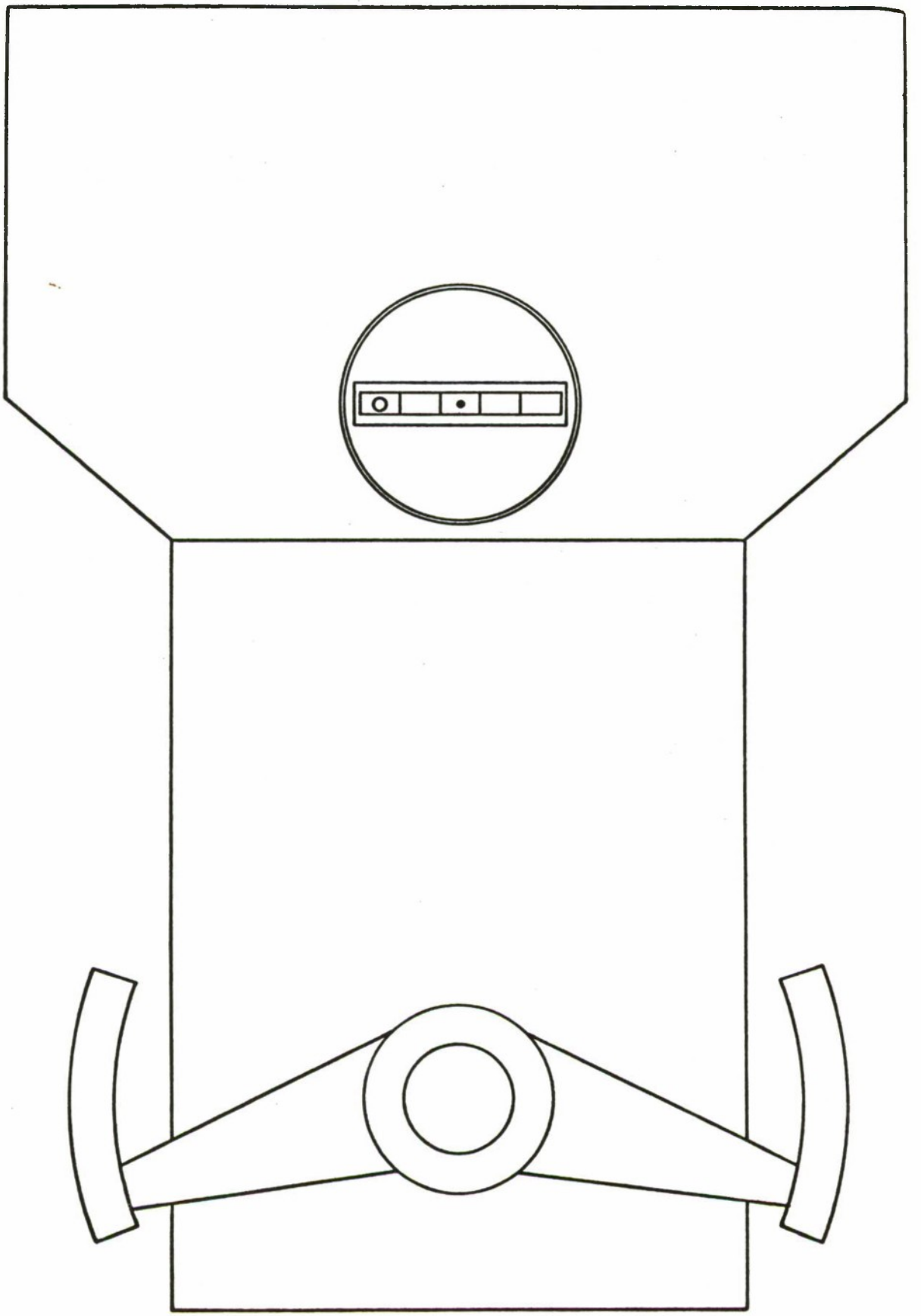
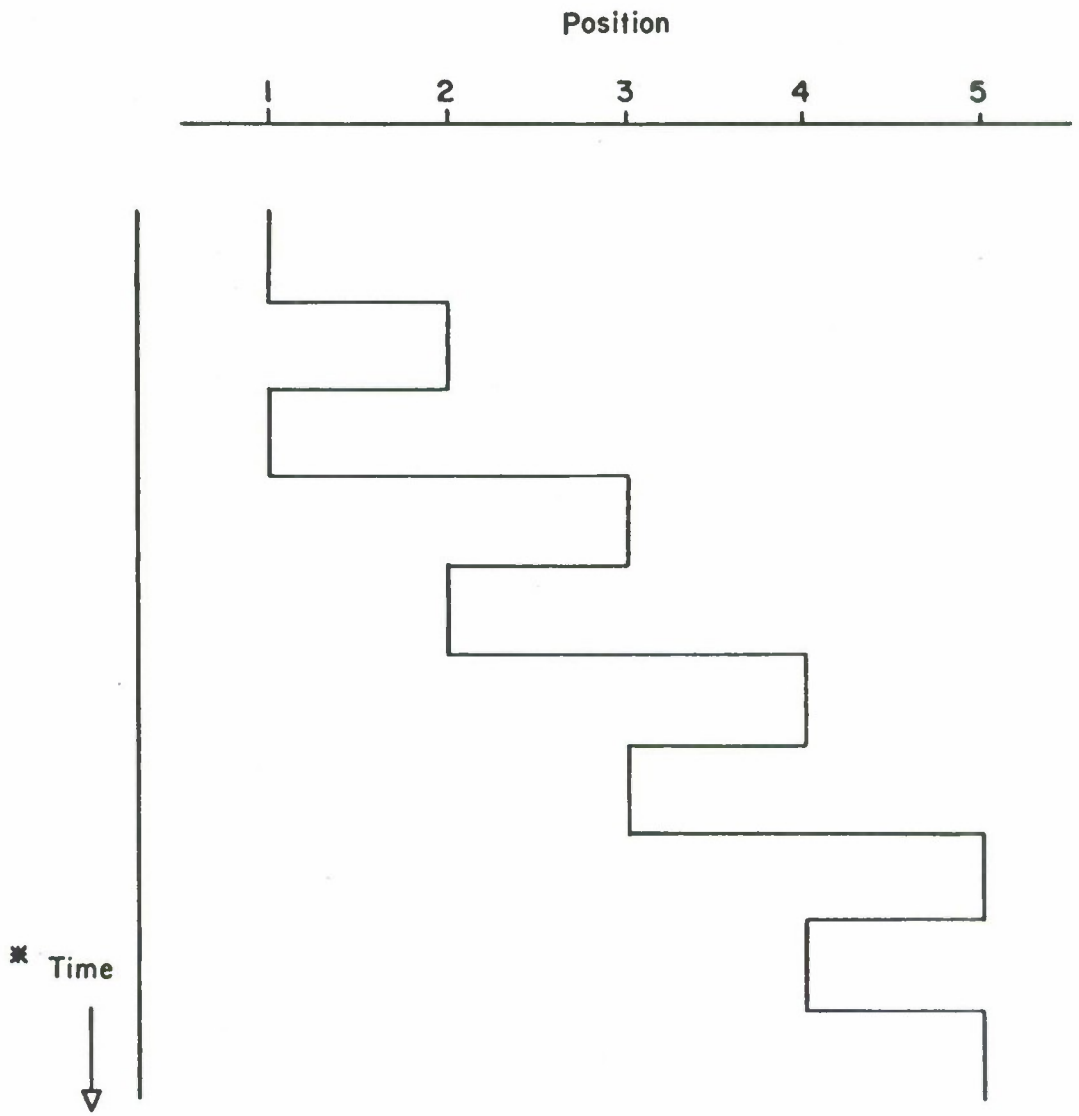


Fig. 1

Target pattern in training runs



* The time scale is under subject control

Fig. 2

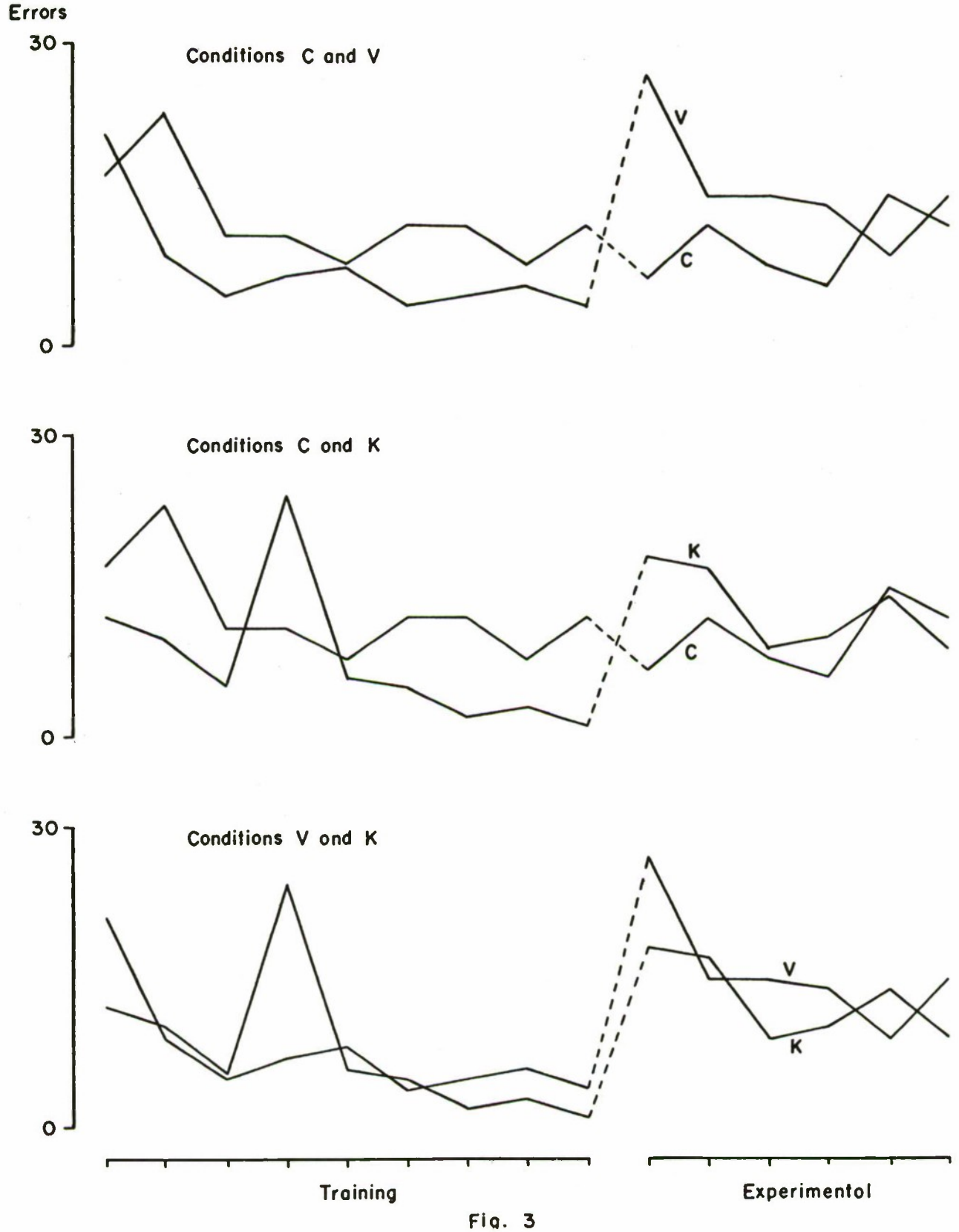

Errors
30
Conditions C and V
V
C
0
30
Conditions C ond K
K
C
0
30
Conditions V ond K
V
K
0
Training
Experimentol

Fig. 3

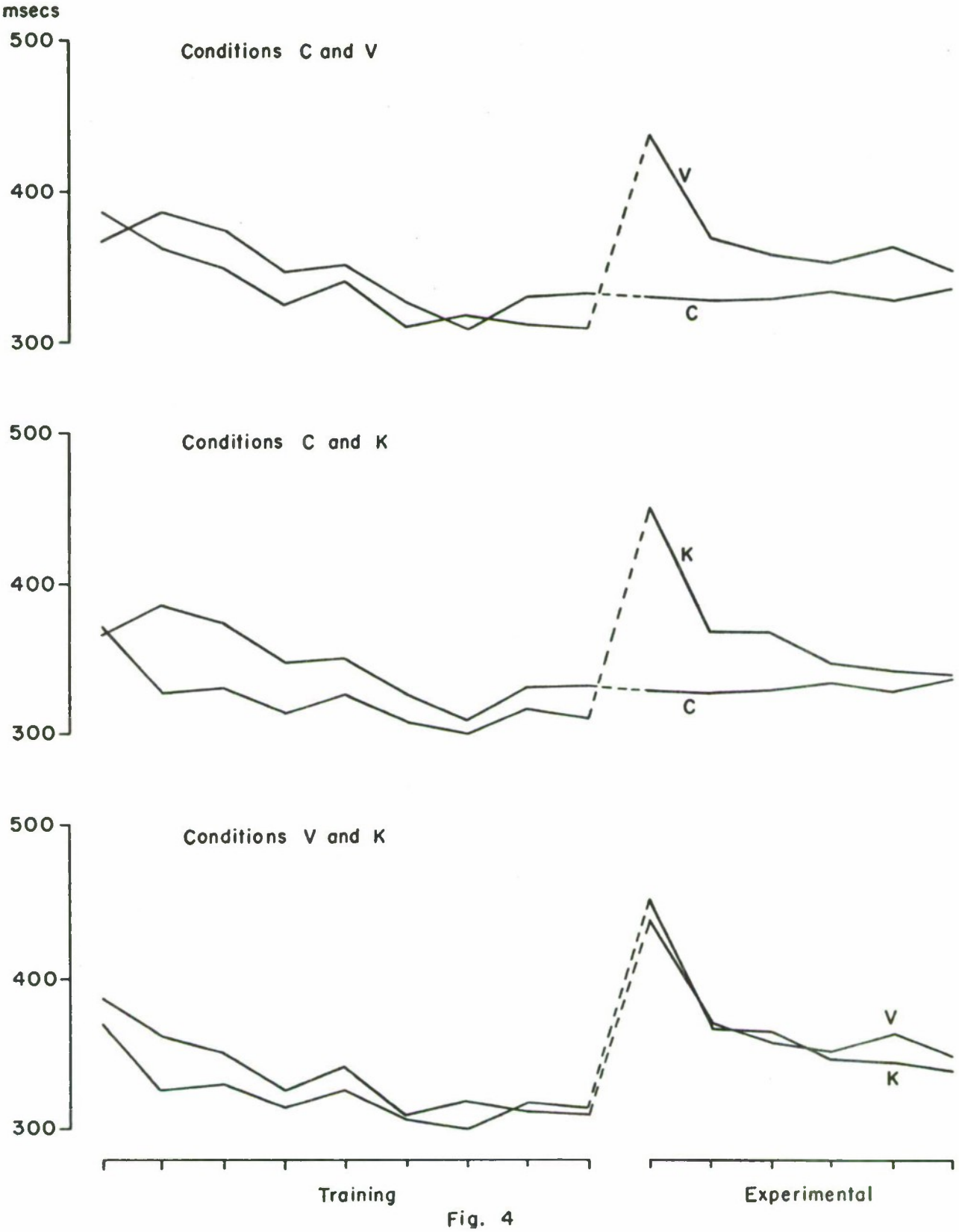

msecs
500
400
300
Conditions C and V
V
C
500
400
300
Conditions C and K
K
C
500
400
300
Conditions V and K
V
K
Training
Experimental

Fig. 4

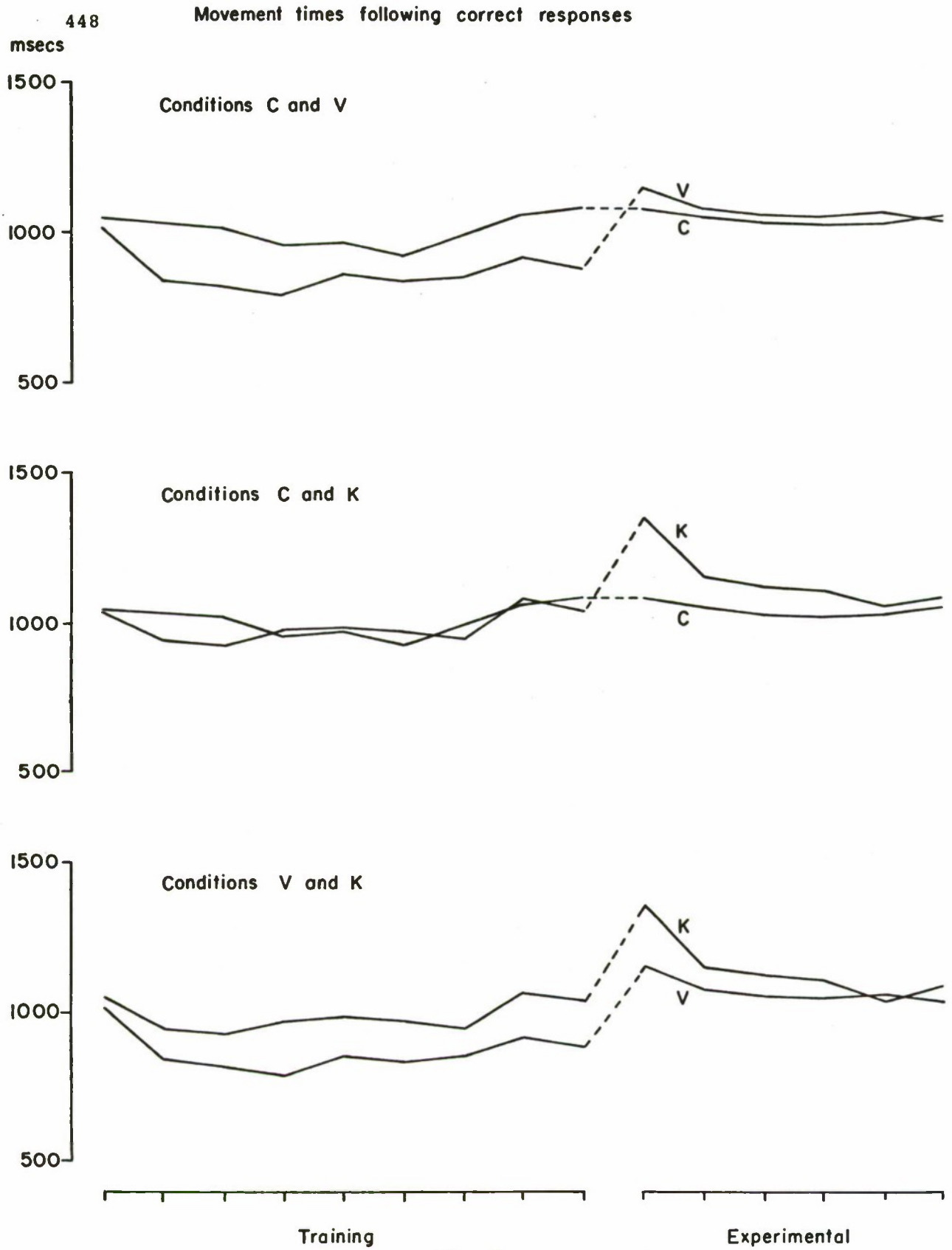

Movement times following correct responses
msecs
1500
1000
500
Conditions C and V
V
C
Conditions C and K
K
C
Conditions V and K
K
V
Training
Experimental

Fig. 5

Training for the Production of Memorized Movement Patterns[1]

Terry R. Armstrong

The University of Michigan
Department of Psychology
Human Performance Center
Ann Arbor, Michigan

Abstract

The effect of augmented proprioceptive feedback information during training for the production of a memorized movement pattern was investigated. The data collected revealed that the guidance techniques used served as aids to performance during training but did not produce performance better than a control during testing. Further studies are underway.

Many skilled control tasks involve the production from memory of movement patterns which require both temporal and spatial accuracy. Common examples might be found in various industrial or vehicle control situations, such as a pilot executing a procedural turn under VFR conditions, or in various sport skills, such as a golf swing or a baseball pitch. For tasks of this type the entire movement has an ideal temporal and spatial pattern, any deviation from this ideal either in extent of movement or in timing will produce unwanted error. In tasks involving movement patterns of this type the operator typically receives no information about how well he is performing until after he completes the entire pattern. Summary error information is provided but no instantaneous error information is available to the performer. The experiments reported here concern the evaluation of various training techniques for the production of memorized movement patterns. Each of the techniques investigated involves providing an operator with some form of instantaneous error information during the training period with the idea that it will not be available when the task is normally carried out.

The human operator may be viewed as an information processing system. Information is received through the senses, is processed by a central decision maker, and is put out through the effectors. There are both external (vision, audition, etc.) and internal (proprioception) systems for monitoring the information output. This feedback, or monitored information concerning the state of the system output with respect to a desired goal, becomes information input to the entire system. It has become clear that information feedback is one of the strongest, most important variables controlling performance and learning. As previously stated, tasks that involve memorized movement patterns typically provide only summary feedback information to an operator. No error information concurrent with the actual

performance of the movement pattern is available. The training techniques being evaluated here involve providing augmented feedback information concerning the output of the motor systems of the arm concurrent with the operator performing the movement pattern. The feedback augmentation makes use of both the visual and proprioceptive senses. It seems reasonable to assume that the availability of such augmented feedback during practice for the production of a memorized movement pattern would result in improved memory for the pattern later when such feedback is not available.

Experiment 1

Five groups of subjects, each assigned to a different experimental condition, were trained to guide a position control stick through a specified brief movement pattern. It was emphasized to each subject that during a testing phase he would be required to produce the pattern from memory. Each group consisted of six randomly chosen right-handed male students from the University of Michigan Human Performance Center paid subject pool.

Subjects sat in an enclosed experimental booth and grasped a position control stick with their right hand, keeping their elbow in a rest over the pivot point of the device. The hand grip and arm rest could be moved through an arc in the horizontal plane. Subjects wore earphones which presented moderate intensity white noise to mask room sounds. A half-second break in the masking noise was provided as a warning to signal the beginning of a trial.

The pattern-to-be-learned lasted 3 seconds and involved three direction reversals of the control device. The maximum movement amplitude measured at the hand grip of the stick was an arc of 24 1/2 inches. A 90 amplifier analog computer and associated components was used to generate the pattern, calculate error scores, and generally control the experiment. A Variable Dynamics Control Stick was used which could operate as a normal position stick but also utilized a torque motor to provide operator reaction forces as specified by functions generated on the analog computer (see Herzog [1969] for details of the device).

Each trial consisted of one complete traversal of the three second movement pattern. After each trial the subject received an integrated-absolute-error (IAE) score for that particular trial, was instructed to return the stick to the starting position, and was informed of the number of the next trial. The inter-trial interval lasted 15-20 seconds.

Fifteen trials for a particular subject constituted a trial block. For every subject, the pattern produced on every fifteenth trial was recorded on an XY plotter. The subject was then informed of his IAE score for that trial and instructed to come out of the test booth where he could compare the XY plot of his latest effort with a plot of the specified pattern. The inter-block interval was typically 2 to 4 minutes.

Days 1 and 2 were each made up of five 15-trial blocks for every subject. Day 3 consisted of four 15-trial blocks of regular practice, and two 15-trial blocks of test trials. For the first block of test trials an IAE score was reported to the subject and an XY plot of the fifteenth trial was made available. IAE scores were not reported to the subject during the last block of test trials.

Subjects practiced under one of five conditions for the 210 practice trials. Group 1 (IAE) served as a control and practiced by trial and error receiving only summary error information. The six subjects in this group were shown a plot of the specified pattern and the approximate movement of the control stick was demonstrated to them. They then attempted to learn the pattern by receiving only an IAE score after each trial and a plot of every fifteenth effort. While these subjects were comparing each fifteenth trial with the specified pattern, the experimenter summarized any trends noticeable in their last trial block. For these subjects the first block of test trials was no different from their regular practice trials. However, they did not receive IAE scores on the last block of test trials.

Group 2 (CRT) practiced the movement as a pursuit tracking task (for further studies of visual aiding techniques see Robb and Pew [1968]). The horizontal movement of a central target dot on a 10 inch cathode ray tube (CRT) display exactly defined the pattern-to-be learned. Via the control stick the subject had complete control of two cursor dots immediately above and below the target on the CRT. By superimposing the cursor dots over the target during a trial, the subject could exactly reproduce the correct pattern with the stick. For these subjects concurrent feedback was available from viewing the positions of the target and cursor in conjunction with manipulation of the control stick. During the two blocks of test trials the CRT was turned off and subjects were required to produce the pattern from memory. During the first block of test trials IAE scores were reported and a plot of the last trial was made as during practice. For the last block of trials the error scores were not given.

In the three remaining conditions the subject felt a steep force gradient whenever he deviated to the edges of a target zone that was centered on the path to be generated. These force gradients which provided concurrent feedback to the subject were produced by the computer-controlled torque motor connected to the control stick. Groups 3, 4, and 5 represented different ways to implement this target zone.

The six subjects in Group 3 (STK) were allowed to deviate a particular amount from the specified pattern at any point during the movement. If this error value was exceeded the subject was subjected to a spring restoring force in the direction that would reduce his deviation. The initial deviation allowed as measured at the hand grip was 1 5/8 inches on either side of the defined pattern making for a target zone of 3 1/4 inches. The zone was eventually decreased to 2 1/4 inches as practice progressed and subjects improved. During practice these subjects essentially were restrained from

making large errors in producing the pattern and received instantaneous error information if they deviated beyond the target zone. For the first fourteen blocks of trials such augmented feedback was always available but during the two blocks of test trials the spring was taken out of the stick dynamics so that the device was essentially free-moving. For the first block of test trials IAE scores were reported and an XY plot of the fifteenth trial was made available as during practice. Error scores were not reported during the last block of test trials.

Practice conditions for Group 4 (DER) were similar to those of Group 3, except that the amount of deviation allowed before the spring restoring forces were introduced varied during the course of a trial. The error value allowed on either side of the specified pattern was reduced by a proportion of the first derivative of the defined signal. In this manner the spring forces were activated sooner when the pattern was changing the most rapidly, allowing less error before feedback was provided to the subject. In this condition the error value allowed on either side of the defined pattern varied from 1 5/8 inches to 1/2 inch during the course of the movement. Test trial conditions paralleled those of Group 3.

The six subjects making up Group 5 (GUI) practiced with a very strong spring gradient and the specified pattern as the target zone. Subjects had their arms guided through the pattern to the capabilities of the device but were instructed to try and anticipate the movement rather than passively let the control stick guide their arms. To the extent they were able to correctly produce the pattern, the less pushing or restraining pressure they would feel from the stick. The subjects in this condition were prevented from ever making substantial errors during practice and further, always had instantaneous error information available to them in the form of felt forces on the stick. Since these subjects did attempt to anticipate the stick movement they were presumably using the same motor systems both during practice and the test period when the stick was free-moving. Test trial conditions paralleled those of Groups 3 and 4.

Figure 1 summarizes the main results of the experiment. Average IAE score for each group is plotted across trial blocks. The data here reveal that the training techniques do serve as aids to performance during the practice period. The learning curve for each of the four experimental groups is well below that of the control. Any comparison of these curves during the practice period is in a sense artificial since Groups 3, 4, and 5 were partially restrained from making substantial error. However, it can be seen that each group demonstrated improved performance as the training progressed. The clear ordering of group average error across practice sessions indicates the relative benefit offered by each training method. Group 3 (STK) consistently performed about 30 percent below that of the control group in terms of average IAE score. Groups 2 and 4 (CRT and DER) each performed about 50 percent below the control while Group 5 (GUI) demonstrated a saving of over 80 percent. It is interesting to note that while Group 5 was restrained from ever making substantial error as evidenced by the average score for trial block 1, they do demonstrate

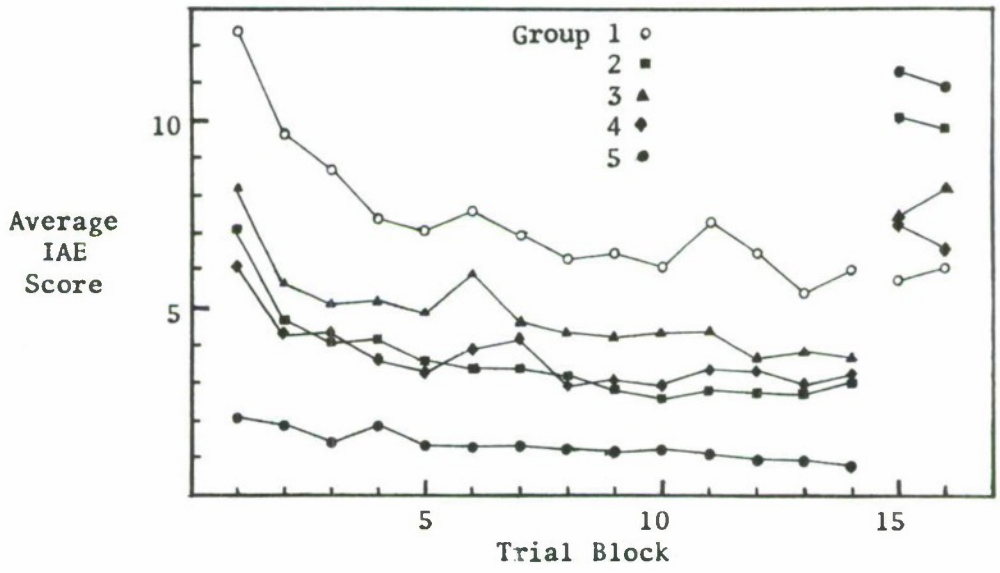


Figure 1. Mean IAE of groups by trial blocks

constant improvement across practice sessions. This can be taken as evidence that these subjects were doing more than passively letting their arms be guided by the control stick. An examination of the XY plots made of each fifteenth trial also substantiates this observation. The pattern produced by the spring-activated control stick and a passive arm in each of conditions 3, 4, and 5 demonstrated some characteristic lags and leads from the ideal pattern. It could be seen that by the second or third block of trials these characteristics were absent in each subject's record and from subsequent records. The experimenter also monitored each trial for every subject visually on a CRT display identical to that used by Group 2 and was able to encourage subjects to actively practice the pattern.

When tested for performance of the memorized pattern, as summarized by trial blocks 15 and 16 on Figure 1, none of the augmented training conditions produced results better than those of the control condition. A three-way analysis of variance of groups x trials x blocks (Winer, 1962) was performed on the individual trial scores making up the averaged test points shown. This analysis revealed that the main effect of groups was significant at the .01 level, the two-way interaction of blocks x trials was significant at the .01 level, and the three-way interaction of groups x blocks x trials was significant at the .05 level. All other effects were not significant. The Newman-Keuls method (Winer, 1962) was used to test the differences between groups on the mean test scores. The results of this test showed that at the .05 significance level Group 2 (CRT)

differs from both Group 1 (IAE) and Group 4 (DER); Group 5 (GUI) differs from Group 1, Group 4, and Group 3 (STK). All other differences between groups were not significant.

The main results of Experiment 1 can be summarized by stating that none of the augmented feedback training techniques produced results better than that of a control group which received only summary error information. Two training conditions, 2 (CRT) and 5 (GUI) produced results significantly poorer than the control. There did appear to be an inverse relation between the amount of aiding provided during training and the ability to reproduce the pattern from memory.

Each of the test trials for every subject was recorded on a strip-chart recorder. A visual examination of these trials failed to reveal any consistent group differences that would help illuminate the main experimental results. A detailed examination of the two test trials recorded by the XY plotter for each subject was made to determine the nature of error. Two types of error could be made, timing error where the subject performed the proper extent movement but at the wrong time, or amplitude error where the extent of the movement was incorrect but the timing was exact. A combination of the two was also possible. The analysis revealed an interesting fact. When timing error was made the entire movement tended to be either compressed or stretched, keeping the proportions of total movement time spent in traversing particular segments of the pattern constant. No such regularity was apparent for amplitude error. It was as though subjects developed a motor program for the pattern which was largely time based. They tended to run the programmed pattern off as a unit keeping the correct timing proportions but unable to keep the total time constant. This finding was observed for subjects in all groups and consistent differences between groups were not apparent.

Several possible factors may have contributed to the negative main result. It may be that for this specific task the detailed summary information provided to the control subjects was exactly compatable with the task requirements and produced the best possible learning. Subjects in the augmented feedback conditions may have been learning how to use the feedback information merely to reduce their error scores rather than attempting to develop a motor plan allowing them to produce the pattern. When they were required to perform without the customary feedback during the test trials, they were unprepared and in a sense found themselves in an entirely new task. This would seem to imply that a particular training method would be task specific.

Several methodological factors also may have been operating. It is possible that the usefulness of augmented feedback varies according to the stage of practice. There is reason to believe, for example, that use of the proprioceptive sense in motor skills becomes important only late in practice (Fleishman and Hempel, 1954). Further, it may be that the limited number of test trials required of each subject could not reveal

differences specific to particular feedback conditions. It is possible that subjects with augmented training could improve during extended testing and surpass the performance of the control group.

Experiment 2

To evaluate some of the factors cited as possible contributors to the main result of the first experiment, further testing of augmented feedback training methods was carried out. Experiment 2, which is now in progress, makes use of four of the five conditions of the first experiment. Groups of eight subjects each, corresponding to the control and conditions 2 (CRT), 3 (STK), and 5 (GUI) were utilized. Subjects were trained to guide the Variable Dynamics Control Stick through a more complex four-second movement pattern to examine the question of task specificity. Throughout three days of practice, subjects in the augmented feedback conditions alternated between practice and test trials so that all subjects were accustomed to the test situation. An entire fourth day was given over to test trials for all subjects. The results of this experiment should help to further illuminate the usefulness of enhanced feedback techniques for training in tasks requiring memorized movement patterns.

Footnotes

[1]This research was supported by the National Aeronautics and Space Administration under Contract NASr-54(06); Dr. R. W. Pew Project Director.

References

Fleishman, E. A., and Hempel, W. E. Changes in factor structure of a complex psychomotor task as a function of practice. Psychometrika, 1954, 19, 239-252.

Herzog, J. H. Proprioceptive cues and their influence on operator performance in manual control. NASA Contractor Report, NASA CR-1248, 1969.

Robb, M., and Pew, R. W. Skill training for the production of a memorized movement pattern. NASA Contractor Report, NASA CR-1251, 1968.

Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw Hill, 1962, pp. 80, 139.

ON SIMULATED AND FULL-SCALE CAR FOLLOWING*

Robert E. Fenton, James H. Ott, and Ronald G. Rule
Department of Electrical Engineering
The Ohio State University
Columbus, Ohio 43210

ABSTRACT

A moving-base automobile simulator was designed and constructed for study of the driver-vehicle interface under steady-state, car-following conditions. It consists of four primary components--a moving-base cockpit, a variable-speed treadmill on which is continuously positioned both a two-lane road and a model lead car, a closed-circuit television system, and a laboratory analog computer. The moving-base feature was included to simulate both the lateral and longitudinal forces inherent in a driving situation.

Describing-function models of an aided driver-vehicle system in a steady-state, car-following situation were obtained using this simulator. A corresponding set of models were obtained under similar full-scale conditions and compared with the simulator set. The models were similar for the two cases; thus the simulator gave a reasonable reproduction of the dynamic behavior of the full-scale, driver-vehicle system. Therefore, it seems probable that other driving situations, which are not too different from the one examined here, can be profitably studied using this simulator.

INTRODUCTION

An automobile simulator is an especially useful device for studying the driver-vehicle system. It enables a researcher to safely conduct a wide variety of tests under closely controlled and easily

---

*This study was sponsored by the Ohio Department of Highways and 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 Ohio or the Bureau of Public Roads.

repeated conditions at a modest cost relative to that of full-scale tests. Further, testing can easily be conducted the year around without regard to the weather. These benefits have resulted in the construction of many different types of driving simulators which have been used to study a wide variety of tasks and driver characteristics. [1-12]

One driving task which has been extensively studied under simulated conditions is steady-state car following. [10-12] Much useful information was obtained from these studies; however, the number of driver characteristics which could be profitably studied were limited by simulator fidelity--particularly the lack of physical motion of the driver's cockpit.

In an effort to overcome both this and other shortcomings, a moving-base, car-following simulator was designed and constructed. Since the simulator was designed for studying various problems associated with driving behavior in a partially automated vehicle, some unusual and innovative instrumentation was incorporated into its structure which is described in the following section.

SIMULATOR DESCRIPTION

The designed simulator consists of four primary components--a moving-base cockpit, a model lead car on a variable-speed treadmill on which is positioned a two-lane road, a closed-circuit television system, and a laboratory analogue computer (See Fig. 1). A qualitative description of these components is presented here and a more detailed technical description is contained in Reference 13.

a) Driver's Cockpit

To simulate the driver's immediate environment, a cockpit with a built-in driver's seat was constructed. A television monitor, simulating the windshield of the driven car, is mounted on the cockpit some 30 inches directly in front of the driver's eyes. An instrument panel containing a speedometer and a gas gauge is located below this monitor.

The cockpit is completely enclosed and mounted on a steel structure so that it can be tilted both fore-and-aft and from side-to-side as shown in Fig. 2. The fore-and-aft (pitch) motion was incorporated to simulate the longitudinal acceleration forces felt by the driver and the side-to-side (roll) motion to simulate the lateral acceleration forces. These motions, which are modelled after an approach used by Hulbert and Wojcik,[4] are accomplished via simple electrohydraulic positioning servomechanisms. The pitch servo is also used to achieve effects analogous to vehicle vibrations and road noise. This is accom-

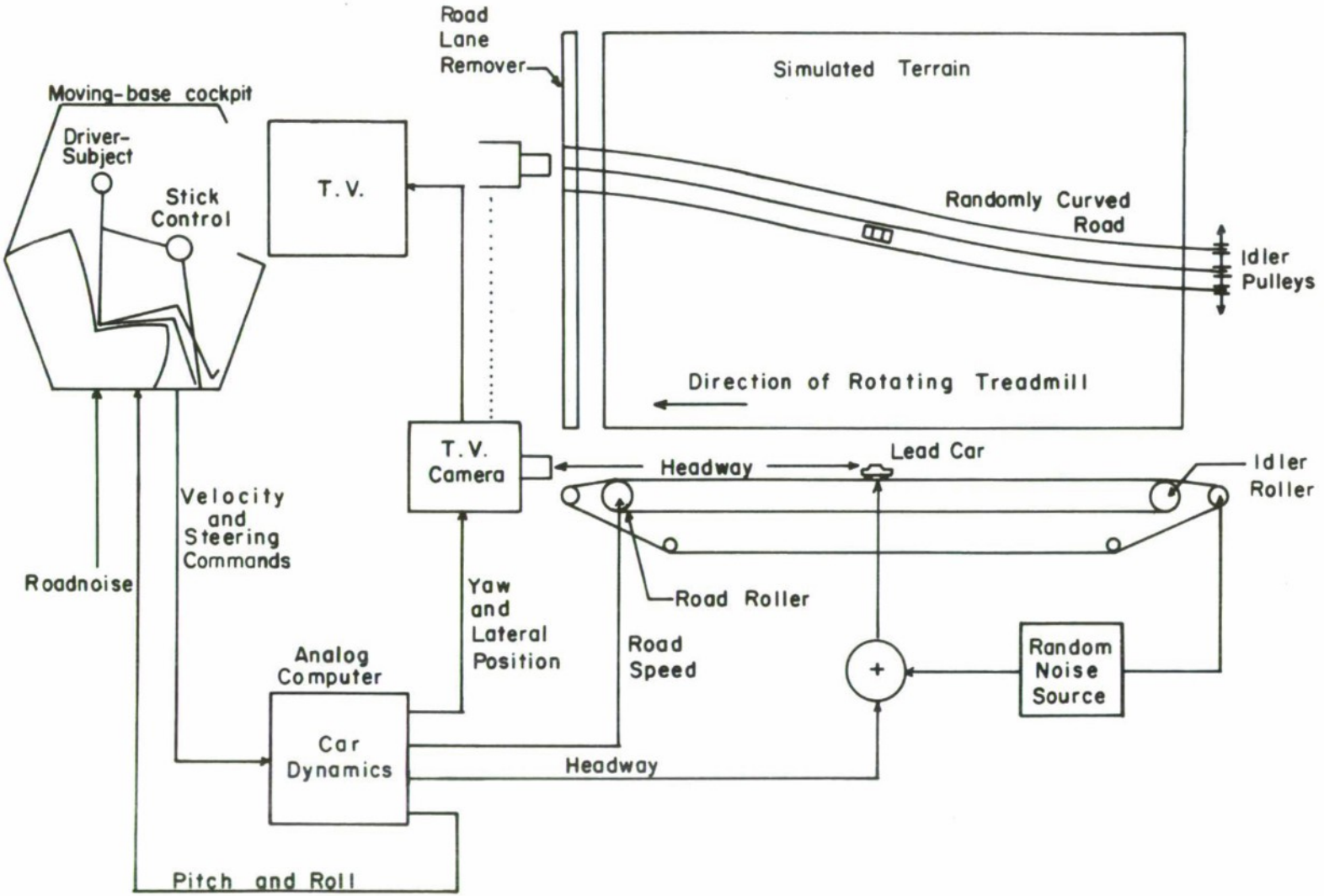


Fig. 1--A moving-base, car following simulator.

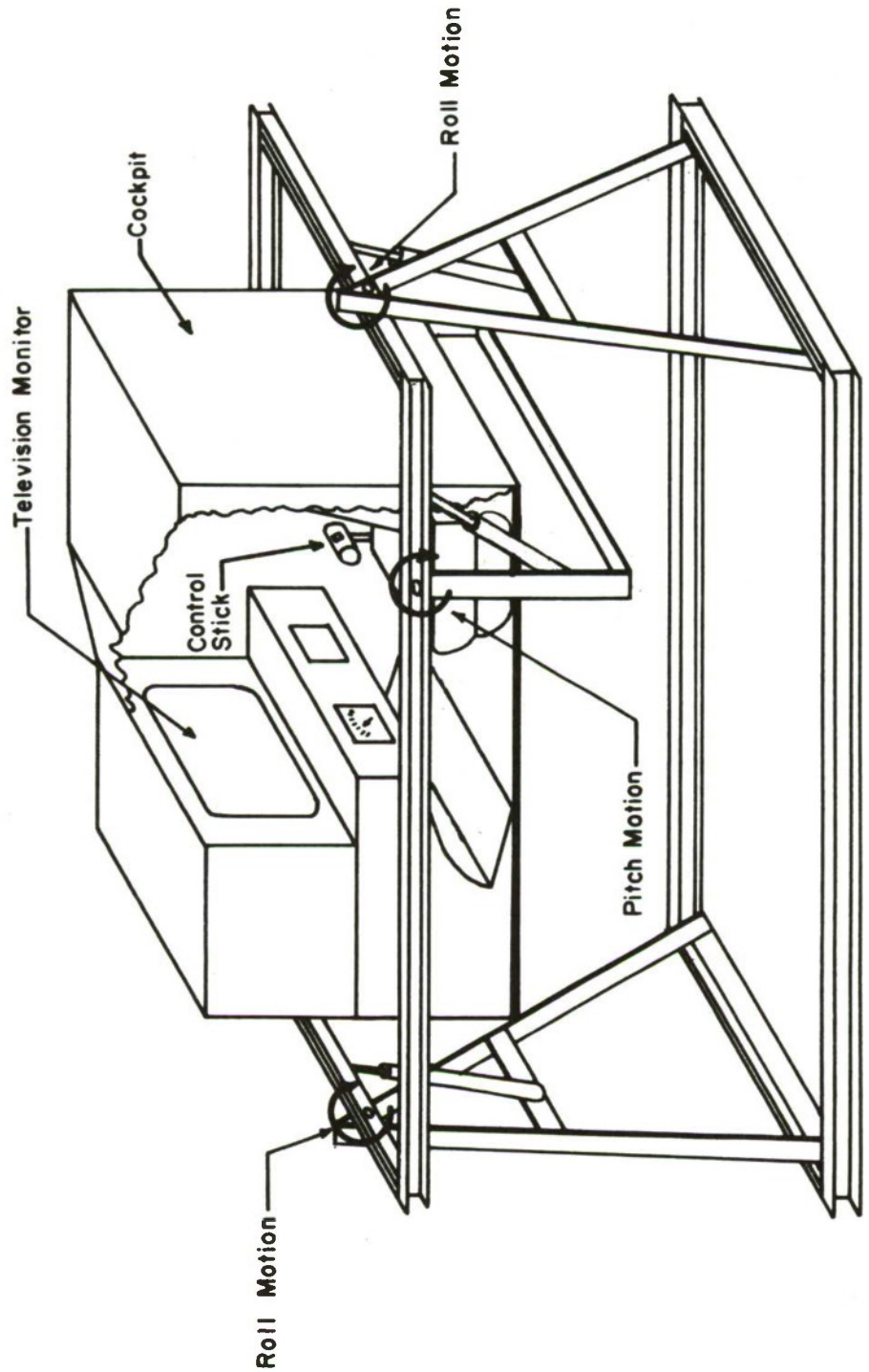


Fig. 2--Moving-base simulator cockpit.

plished by exciting this servo with a random signal which causes small random displacements in pitch angle.

The cockpit is equipped with a control stick in lieu of conventional automobile controls. This was done so that data could be obtained which would be useful in future studies involving a partially automated vehicle which contains a similar control stick.

A speaker is included in the cockpit to present audible sound, approximating car and wind noise, to the driver-subjects.

b) Model Roadway

The model roadway consists of a treadmill on which is positioned a two-lane road and a miniature lead car. The lane strips which define the road are "pressed" onto the treadmill belt at the right end and "peeled" off at the left end (See Fig. 1). Many roadway curvatures can thus be generated while the treadmill is rotating. The lead car is a scale model (87:1) which is kept between the lane lines by a simple steerable mechanical follower. The effective headway between this vehicle and the controlled one is continuously adjusted to the proper value by a simple position servomechanism.

c) Closed-Circuit Television System

A high resolution (800 lines) closed-circuit television system is used to display the two-lane roadway and lead car to a subject as shown in Fig. 1. The camera is moved laterally to simulate lateral position of the vehicle, and it is rotated to simulate both the yaw rate and the side-slip angle of the vehicle.

The video monitor in the cockpit receives the output signal from this camera and the resulting visual display is a picture of the lead car and two-lane roadway against a neutral background. The primary visual cues available to a subject are those associated with the observed relative motion of the model lead car; however, absolute speed information is also available from the speedometer mounted under the monitor.

d) Analog Computer and System Dynamics

A laboratory analog computer is used to simulate the automobile dynamics and to compute various measures of driving performance.

The vehicle longitudinal dynamics were chosen so as to approximate those of a 1965 Plymouth sedan and are

$$\frac{e^{-0.17s}}{20s + 1} \qquad (s = \sigma + j\omega). \qquad (1)$$

The lateral dynamics were modelled using the following set of equations which were developed by Segal:[14]

$$\frac{\dot{\alpha}v_2}{100} = -\frac{\dot{\psi}v_2}{100} - 0.011\ddot{\theta} - 2.5\alpha - \frac{0.87\dot{\psi}}{v} + 0.562\theta + 1.2\delta_s \qquad (2)$$

$$\frac{\ddot{\theta}}{10} = -\frac{\dot{\alpha}v_2}{50} - \frac{\dot{\psi}v_2}{50} - 0.07\ddot{\psi} - 0.45\dot{\theta} - 4.25\theta \qquad (3)$$

$$\frac{\ddot{\psi}}{10} = -0.13\ddot{\theta} - 0.18\alpha - \frac{35.2\dot{\psi}}{v_2} + 1.08\theta + 2.45\alpha_s + 0.002\dot{\theta} \qquad (4)$$

where α = side-slip angle
ψ = yaw rate
θ = roll angle
δ_s = steering command input
v_2 = speed of vehicle

If the vehicle speed is constant, these equations are linear. Various response curves obtained from a computer solution of the linearized equations closely approximated those obtained from road tests of a 1965 Plymouth sedan.

EXPERIMENTAL STUDIES

The results obtained using an automobile simulator should approximate those obtained from a corresponding study in a highway situation. If close correlation were obtained, one would have considerable confidence in the real-world applicability of results obtained from future similar studies on that simulator.

An abreviated simulator validation study, which consisted of two experiments in driver-aided car-following, is discussed in this section. The first experiment was conducted on the moving-base simulator, and a second corresponding one was conducted using a test vehicle on the highway.

a) A Driver-Aid System

An aided driver-vehicle system can be represented by the block diagram shown in Fig. 3. This model is intended to correspond to the case of steady-state car following on a long, straight and uncrowded superhighway; thus it is assumed here that the driver is primarily concerned with longitudinal control and that only small lateral corrections need be made. The system input is the velocity of a lead car $v_1(t)$, and the output $v_2(t)$ is the velocity of the controlled one. Note from the figure that the driver receives an input from a display and controls the vehicle via a control stick.

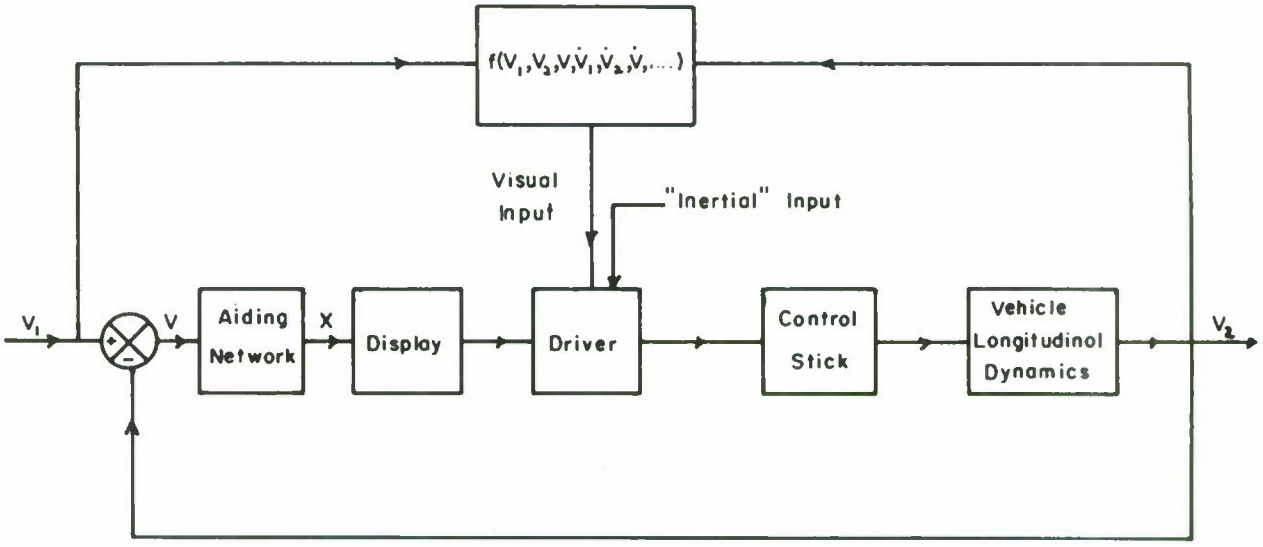


Fig. 3--Block diagram of driver-aided system.

The control stick, which is shown in Fig. 4, is located on the right side of the driver and replaced the steering wheel, brake pedal, and accelerator pedal. The controlled vehicle is steered by moving the control stick head to the left or right, accelerated by moving it forward and braked by pulling it back. The control stick design depicted in Fig. 4 was chosen so as to eliminate cross-coupling between the lateral and longitudinal modes which has previously been encountered with a front-mounted, fully rotational control stick.[17]

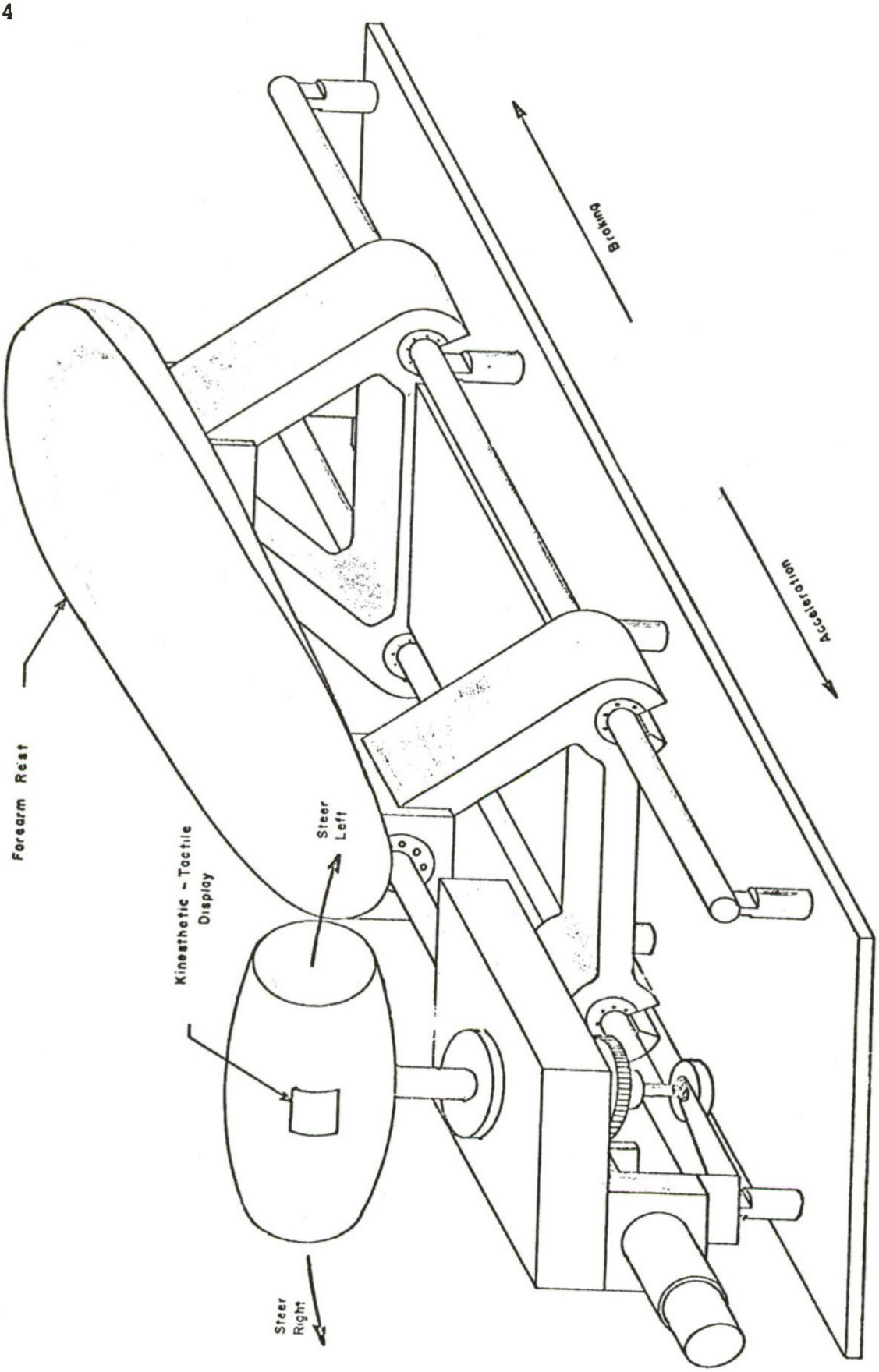


Fig. 4--Side-mounted control stick with built-in kinesthetic-tactile display.

A kinesthetic-tactile, compensatory display was built into control stick head as shown in Fig. 4. A subject determines the magnitude and direction of an error by feeling the displacement of a metal finger from the neutral or "flush" position with respect to the control stick head. When the finger protrudes, he moves the stick forward, and when the finger recesses he pulls the stick back; thus, the subject "follows" the finger and attempts to keep it flush with the control stick head.

The position of the finger is controlled by a servomechanism with a speed of response less than 0.2 second. The servo input signal (x) is the output of the aiding network shown in Fig. 3. This network consists of a circuit whose output is a linear sum of two terms--one proportional to the network input and the second to the integral of the input. Since the network input is relative velocity (v)

$$x = (k_1 \Delta h + k_2 v) k \qquad (5)$$

Here, k_1 and k_2 are constants, k is a gain term and Δh is the deviation from some desired headway between the lead and controlled vehicles. Since the response of the display servo is quite fast, the signal displayed to the driver is defined by Eqn. (5).

The constant k_1 in Eqn. (5) determines the finger-position sensitivity, which refers to the finger movement per unit headway deviation. The mathematical ratio (k_2/k_1) defines the term aiding ratio, and it provides a relative measure of how much derivative information is supplied to a subject. Since relative velocity is the derivative of headway, the servo input signal drives the finger in a manner that generates a partially quickened, compensatory display. In previous highway testing of this driver-aid system, it was found that excellent results were obtained for $k_1 = 0.064$ in/ft and $k_2/k_1 = 1.5$.[12]

In addition to the display input shown in Fig. 3, a driver also has his usual visual input and one from motion cues. If the driver controls the vehicle almost solely on the basis of his "normal" visual information input and the display input, then the overall system can 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.[15] This model will be of the form

$$G(j\omega) = \frac{\phi_{v_1 v_2}(j\omega)}{\phi_{v_1 v_1}(j\omega)} \qquad (6)$$

where

$\phi_{v_1 v_2}(j\omega)$ = cross-power spectrum of lead and driven car speeds,
$\phi_{v_1 v_1}(j\omega)$ = power spectrum of lead-car speed.

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 (6) can also be written as

$$G(j\omega) = \frac{V_2(j\omega)}{V_1(j\omega)} \tag{7}$$

where $V_1(j\omega)$ and $V_2(j\omega)$ are the Fourier transforms of those parts of $v_1(t)$ and $v_2(t)$ which are linearly correlated.

In this study such linear dynamic models were obtained from both a simulator study and a corresponding real-world study and the models compared for purposes of simulator validation.

b) Simulator Testing

In the first phase of the testing, a study of driver-aided car following was made on the automobile simulator. Subjects were tested in a steady-state car-following situation, and quantitative models of the resulting driver-simulator system were obtained.

i) Experimental Description

The lead-car velocity consisted of an average value of 40 mph plus a low-frequency random component. The corresponding accelerations of the lead car were in the range

$$-0.04\ g \leq \alpha \leq 0.04\ g$$

which approximates the dispersion values of $\pm$ 0.03 g which Montroll[16] has reported as typical for normal car following. The frequency spectrum associated with the lead-car speed was almost identical with that obtained from later highway testing.

The model roadway was randomly curved with the minimum radius of curvature used being 1500 ft.

Each subject was instructed to respond to the display so as to maintain an invariant headway between himself and the lead car. Hence, his task was car-following on a randomly curving road.

Two male subjects aged 19 and 25 were used in this test. Both had extensive driving experience using the side-mounted control stick and the kinesthetic-tactile display. Their experience was gained on both the automobile simulator and an instrumented vehicle.

Two 5-minute runs were made for each subject. A subject was first shown the lead car at a simulated headway of 40 ft.* and told to maintain that headway. Approximately one minute after he assumed control, a magnetic-tape recorder was started and the variational components of both $v_1(t)$ and $v_2(t)$ recorded for the next 5-minutes. These signals were later processed in order to obtain a system model per Eqn. (6).

These simulator tests and the real-world ones to be described later were conducted over a three-day period. The subjects drove the simulator each morning, and the vehicle each afternoon.

ii) Experimental Results

The results are shown in the Bode plot in Fig. 5 where $20 \log_{10} \left| \frac{V_2(j\omega)}{V_1(j\omega)} \right|$ is plotted versus frequency. Four curves, corresponding to two runs each by two subjects are shown. The corresponding phase-angle plots are also shown in this figure.

Note the similarity of the system functions, both within and across subjects, up to approximately 0.8 radian which is the frequency where the estimated linear coherence function rapidly decreased. Thus, it seems clear that the subjects' behaved comparably and consistently.

These data were collected using an effective vehicle time lag of 0.17 sec per Eqn. (1). A similar set of data were collected for an effective lag of 0.25 sec and are contained in Appendix A. A comparison of the Bode plots obtained for the two cases indicates that similar results were obtained for each value of lag.

---

*A headway of 40 ft. corresponded to a bumper-to-bumper spacing of 33 ft.

c) Highway Testing

A fully instrumented 1965 Plymouth sedan was used in the second phase of the testing. The instrumentation included a side-mounted control stick with a built-in kinesthetic-tactile display and a mechanical take-up reel, or "yo-yo", for obtaining headway and relative velocity data. The same subjects who participated in the simulator experiment were tested in a steady-state, car-following situation and describing-function models of the driver-vehicle system were obtained per Eqn. (6). A 1969 Plymouth sedan was used as the lead car.

A 7-mile length of nearly straight and level interstate highway was used for the experimental runs. This highway was open to the general public; however, since it was a completed section of an unfinished highway, public usage was negligible.

i) Experimental Description

The controlled vehicle and the lead car were initially at standstill and separated by the desired spacing of 33 ft. The display was activated and the lead-car driver was instructed to accelerate to, and maintain a speed of 40 mph (His "steady-state" speed actually varied randomly between 37 and 43 mph). The subjects were told to respond to the display so as to maintain a constant headway both during the acceleration phase and the car-following phase.

When both cars were traveling at the same speed at the desired headway, a tape recorder was started and the variational components of both $v_1(t)$ and $v_2(t)$ were recorded for the following 5 minutes. The corresponding system model (Eqn. (6)) was later obtained from these data.

ii) Experimental Results

The results are shown in the Bode plot in Fig. 6 where $20 \log_{10} \left| \frac{V_2(j\omega)}{V_1(j\omega)} \right|$ is plotted versus frequency. Four curves, corresponding to two runs each by two subjects, are shown. The corresponding phase-angle plots are also shown in this figure.

In each case the system response was nearly zero decibels until approximately 0.2 rad/sec; thereafter it essentially rose to a resonant peak near 1 rad/sec and then decreased toward zero decibels as the frequency approached 1.5 rad/sec. The phase angle in each case was nearly zero until 0.5 rad/sec; thereafter it decreased sharply to about

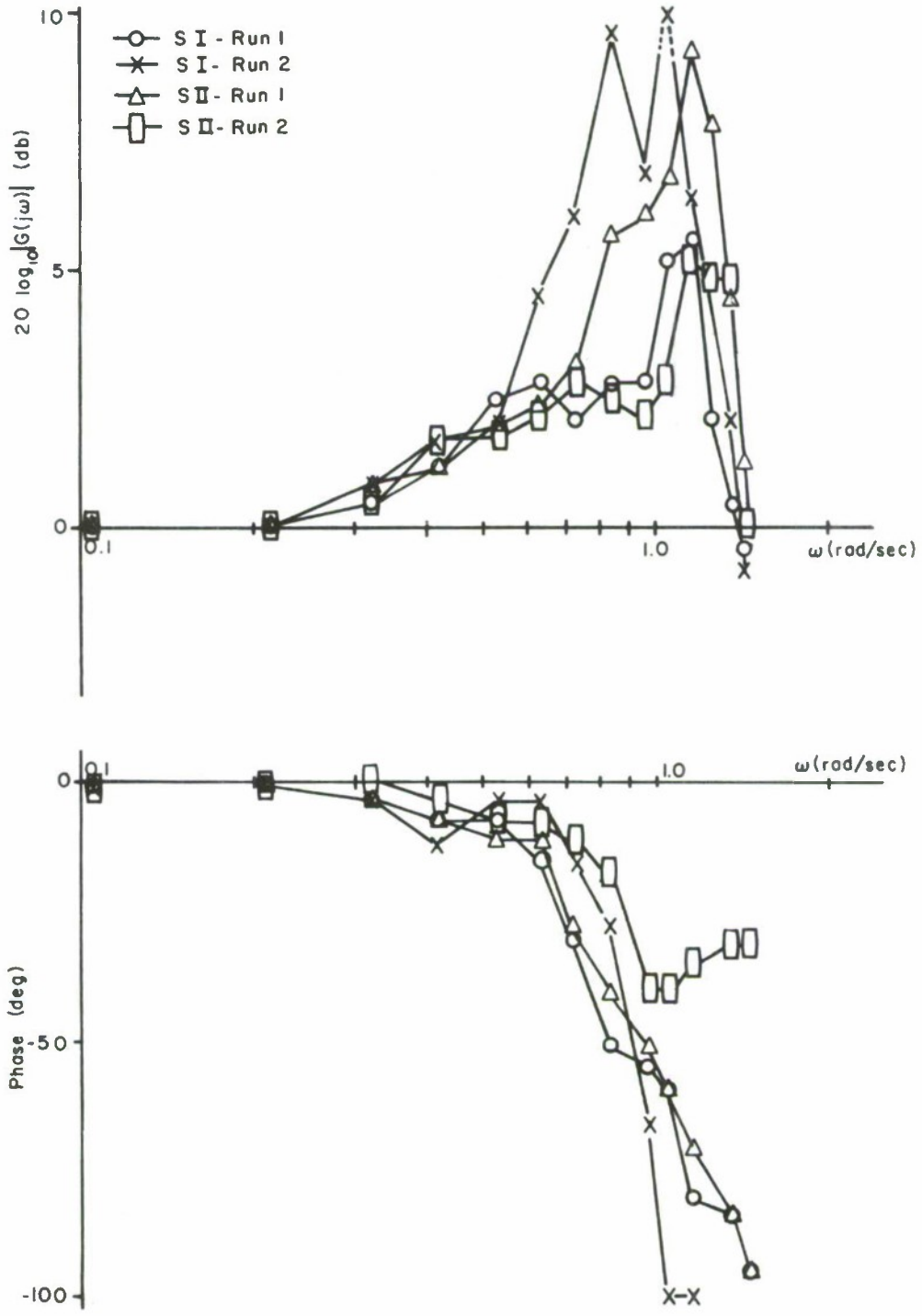


Fig. 5 --Describing-function models of driver-aided system obtained from simulator testing ($\tau = 0.17$).

-90 degrees. The similarity present in these models leads one to conclude that the subjects' behavior was comparable and consistent in car following. It is interesting to note that these models are nearly identical to those obtained previously from four different subjects in a similar experiment in aided car following.[17]

SIMULATOR VALIDATION

In order to simplify the comparison between the real-world results and the simulator results, the data for each class were averaged within and across subjects. An examination of the two averaged magnitude plots and the corresponding averaged phase-angle plots depicted in Fig. 7 shows that essentially the same dynamic model was obtained in each case. This same conclusion is also clear from a comparison of both the individual magnitude plots and the individual phase-angle plots which are all shown in Figs. 5 and 6.*

Thus the longitudinal performance of the simulated driver-vehicle system and the actual driver-vehicle system was essentially the same. It can be inferred from this that the subject-drivers behaved in a similar manner in the two cases; however, this is conjectural at the present time and must be examined in more detail by future experimentation.

CONCLUSIONS

The results reported here were obtained from a limited number of subjects under highly restrictive conditions and should be treated with caution; however, they do suggest the feasibility and desirability of using the simulator in the study of various steady-state freeway driving situations.

It should be noted that although the primary purpose in constructing this simulator is the study of the driver-vehicle system in an automated or partially automated highway environment, it would also be useful in the study of the conventional driver-vehicle system. Thus, such research as a study of driver reaction times under various highway conditions,

---

*The simulator data described in Appendix A were also averaged and plotted in Fig. 7. The resulting average curve is quite similar to the other two averaged curves thus indicating the system insensitivity to small changes in the effective vehicle lag. In essence, all simulator runs can be viewed as being obtained under identical conditions.

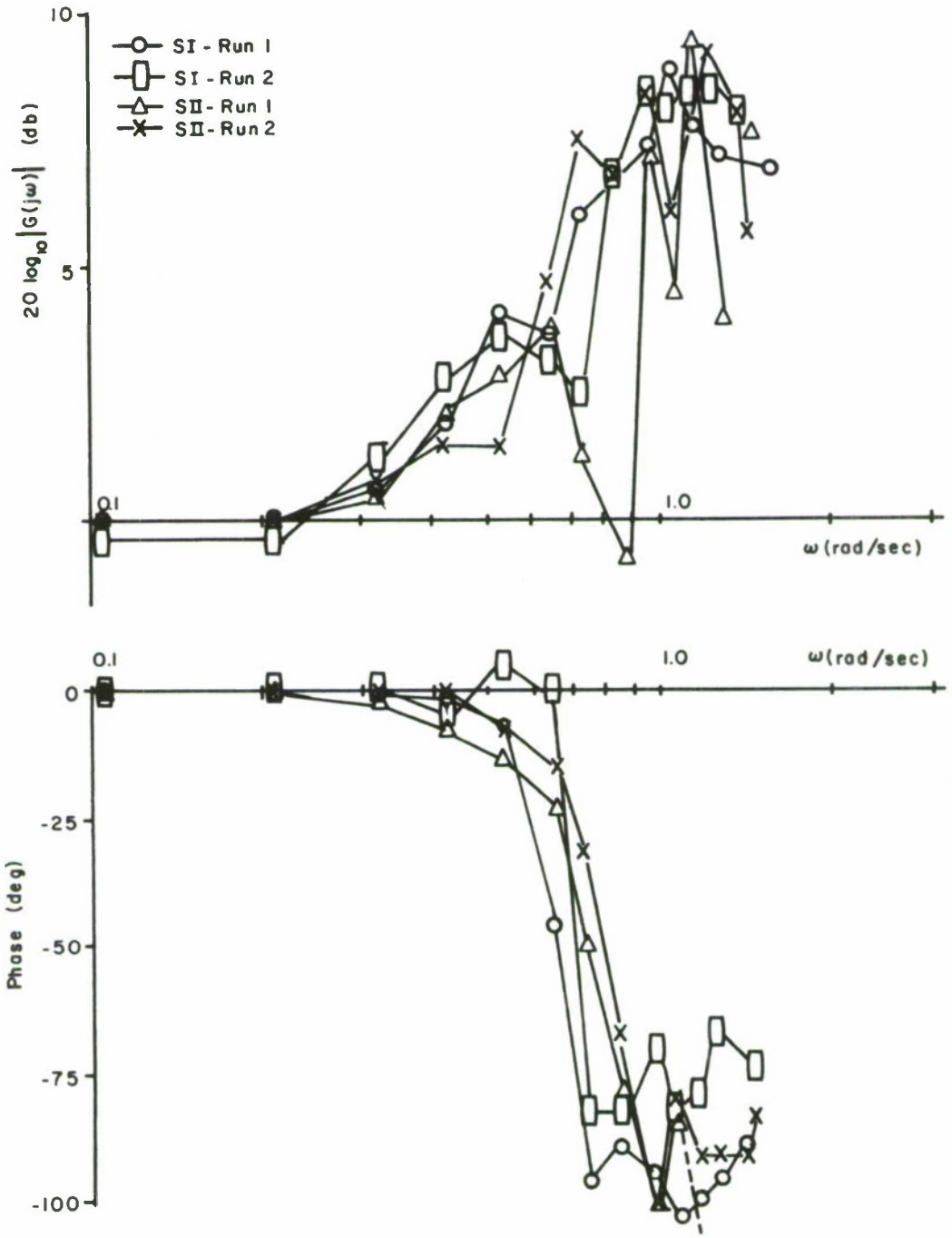


Fig. 6--Describing-function models of driver-aided system obtained from real-world testing.

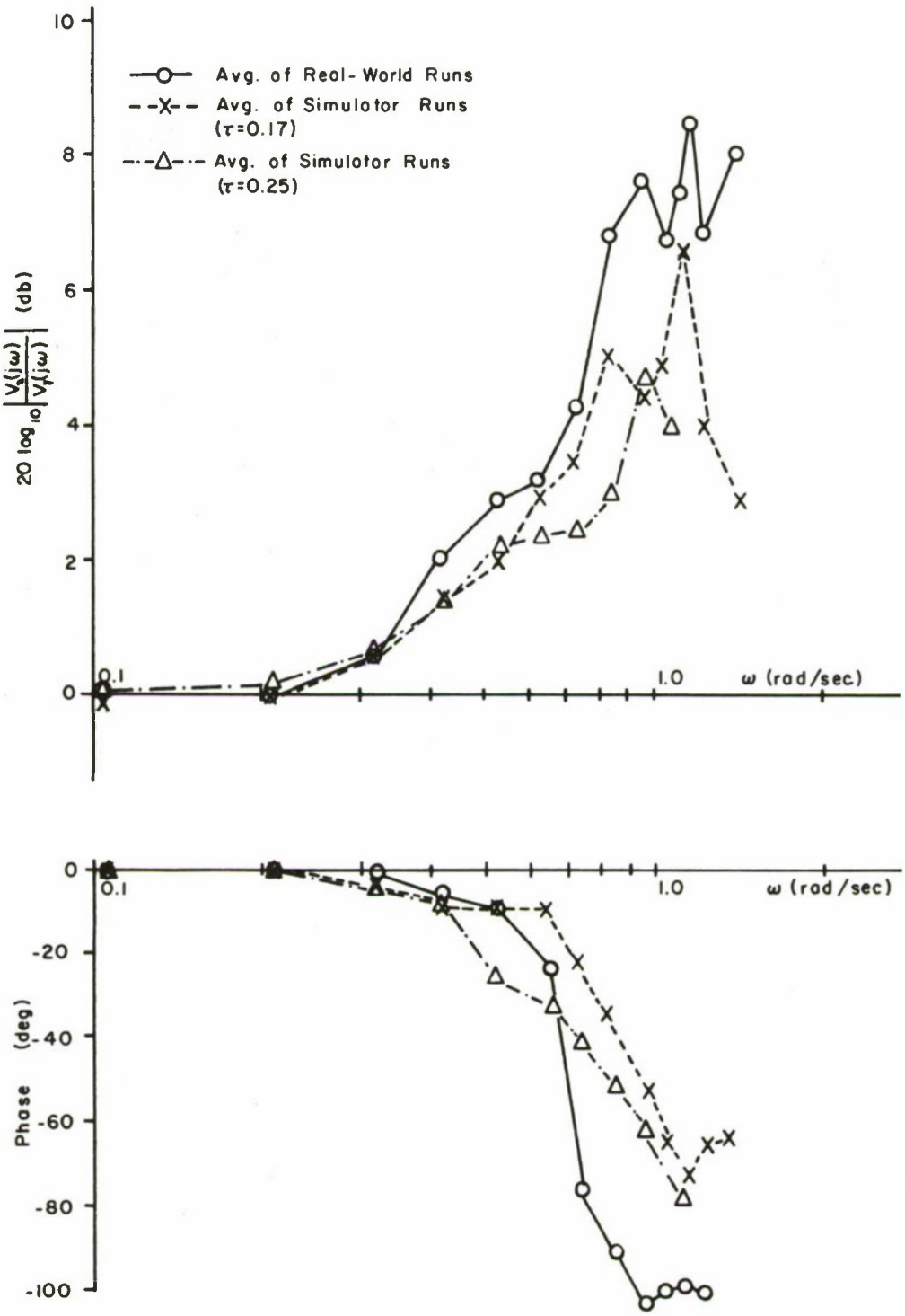


Fig. 7--Average models from both simulator and real-world testing.

the effects of various stimulants and depressants on driving behavior, and driving fatigue could also be efficiently examined within the safe confines of a laboratory.

APPENDIX A

ADDITIONAL SIMULATOR TEST RESULTS

The simulator tests discussed earlier were repeated under nearly identical conditions one day after the first tests were conducted. The single difference in these latter tests was in the choice of vehicle time lag per Eqn. (1). Previously, this quantity was chosen as 0.17 sec as indicated in Eqn. (1), and in the tests reported here, it was taken as 0.25 sec. These values were chosen since experimental measurements of this quantity ranged between 0.17 - 0.25 sec for the instrumented sedan used in the corresponding full-scale car-following studies.

The results obtained from 2 runs on each of the same two subjects who participated in the previous experiments are shown in Fig. 7. Here the closed-loop system response, $20 \log_{10} \left| \frac{V_2(j\omega)}{V_1(j\omega)} \right|$ is plotted versus frequency for each of the 4 runs. The corresponding phase-angle plots are also shown in this figure.

Note the similarity of the system functions--both within and across subjects. This is a strong indication that the subjects behaved comparably and consistently.

A comparison of these data with those collected for a vehicle time lag of 0.25 sec indicates that essentially similar results were obtained for each value of lag (See Fig. 7). In essence, both sets of simulator runs can be viewed as obtained under identical conditions.

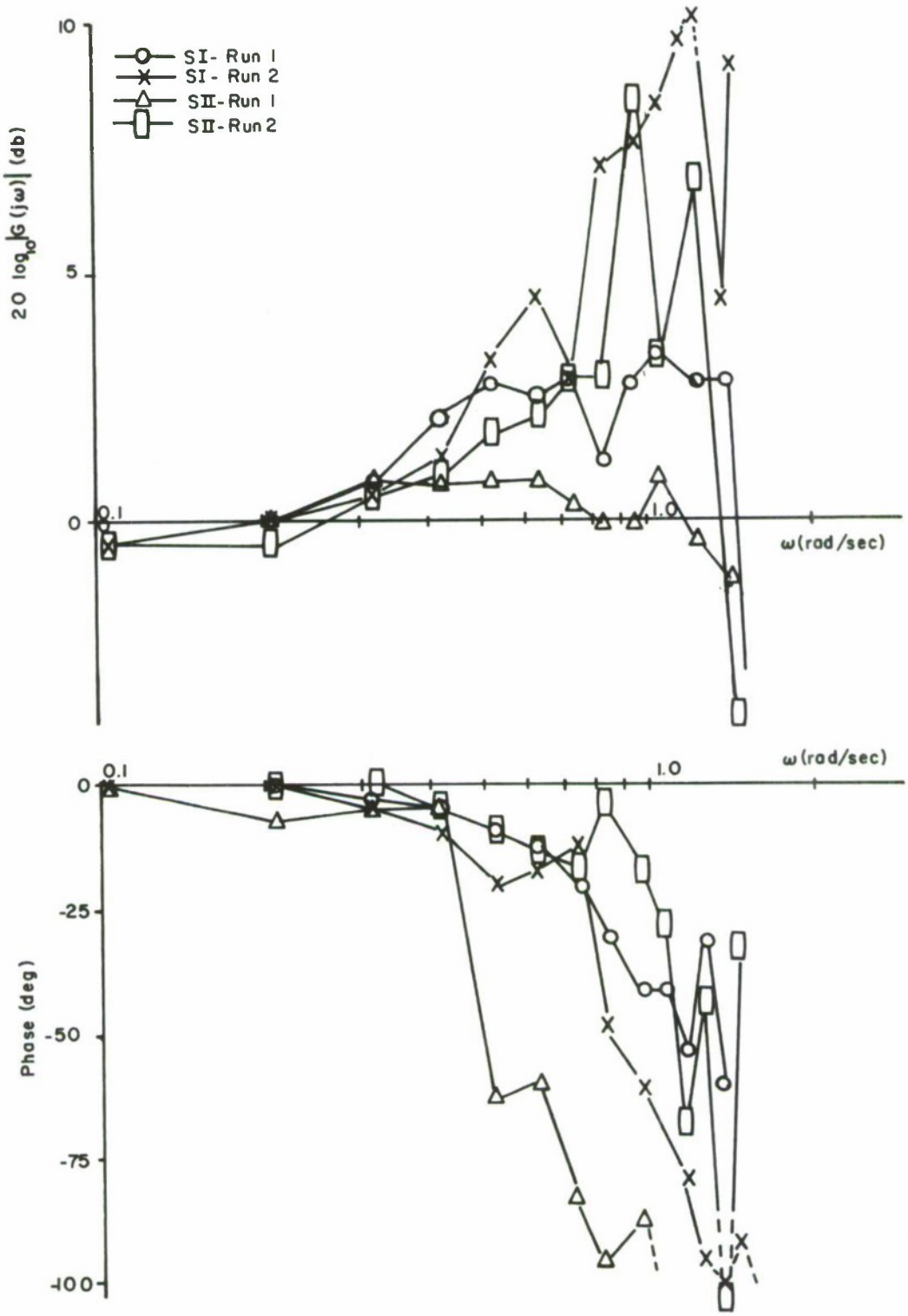


Fig. 8--Describing-function models of driver-aided system obtained from simulator testing ($\tau = 0.25$).

REFERENCES

1. Bidwell, J. B., "Vehicle Control Simulation", General Motors Research Laboratories, Warren, Michigan, Report GMR 319, March 1961.

2. Wierwille, W. W., Gagne', G. A., and Knight, J. R., "An Experimental Study of Human Operator Models and Closed-Loop Analysis Methods for High-Speed Automobile Driving", IEEE Transactions on Human Factors in Electronics, Vol. HFE-8, No. 3, pp. 187-201.

3. Hulbert, S. F., and Wojcik, C. K., "Part Task Simulator", Report published by the Department of Engineering, University of California at Los Angeles, Los Angeles, California.

4. Hulbert, S. F., and Wojcik, C. K., "Transportation Human Factors: Part 1, Physical Forces", Report No. 64-24, Department of Engineering, University of California at Los Angeles, Los Angeles, California, May 1964.

5. Policny, J., "Driving Simulator", Popular Electronics, January 25, 1965, pp. 102-105.

6. Roland, R. D., and Sheridan, T. B., "A Normative Model for Control of Vehicle Trajectories in an Emergency Maneuver", presented at the Annual Meeting of the Highway Research Board, Washington, D. C., 17-21 January 1966.

7. Todosiev, E. P., and Fenton, R. E., "Velocity Thresholds in Car Following at Night", Highway Research Record No. 122, 1966.

8. Todosiev, E. P., "An Automobile Simulator", presented at the 1962 AIEE Empire Tri-District Meeting, Erie, Pennsylvania, AIEE Paper No. DP 62-967.

9. Sheridan, T. B., (Editor), "Proceedings of a Conference on Mathematical Models and Simulation of Automobile Driving", Massachusetts Institute of Technology, Cambridge, Massachusetts, 28-29 September 1967.

10. Fenton, R. E., "An Improved Man-Machine Interface for the Driver-Vehicle System", IEEE Transactions on Human Factors in Electronics, Vol. HFE-7, No. 4, December 1966, pp. 150-157.

11. Fenton, R. E., "Asymptotic Stability Studies in Simulated Car Following", Proceedings of Third Annual NASA-University Conference on Manual Control, National Aeronautics and Space Administration, Washington, D. C., NASA SP-144, 1967, pp. 399-414.

12. Montano, W. B., and Fenton, R. E., "An Investigation of Driver-Aided Car Following", Communication and Control Systems Laboratory, The Ohio State University, Report EES 276A-2, November 1967.

13. Ott, J. H., Shah, F. B. S., Fenton, R. E., and Chaudhry, V. K., "A Moving-Base Car-Following Automobile Simulator", in "The Driver-Automatic System Interface", Communication and Control Systems Laboratory, The Ohio State University, Columbus, Ohio, Report EES 276A-8, April 1969.

14. Segal, L., "Theoretical Prediction and Experimental Substantiation of the Response of the Automobile to Steering Control", Proceedings of the Automobile Division, The Institution of Mechanical Engineers, No. 7, 1956-1957, pp. 310-330.

15. Rule, R. G., and Fenton, R. E., "Modelling the Driver-Vehicle System--A Time-Series Analysis Approach", Communication and Control Systems Laboratory, The Ohio State University, Columbus, Ohio, Report EES 276A-1, March 1969.

16. Montroll, E. W., "Acceleration Noise and Clustering Tendency of Vehicular Traffic", in Theory of Traffic Flow, R. Herman, Editor, Amsterdam: Elsevier Publishing Company, 1961, pp. 147-157.

17. Fenton, R. E., and Montano, W. B., "A Study of Driver-Aided Car Following" presented at the 4th International Symposium on the Theory of Traffic Flow, Karlsruhe, Germany, June 18-20, 1968 (to be published in the proceedings of that conference).

IN-FLIGHT EVALUATION OF PILOT'S CONTROLLERS

by

DAVID E. FREARSON
and
JAMES A. TOWNSEND
Control Systems Research Branch
Flight Control Division
AF Flight Dynamics Laboratory

Current programs for the development of new flight control systems indicate the need for a concurrent exploratory evaluation of new concepts in hand controls (pilot's controllers) to provide system performance capability and acceptance by pilots for application in next generation aircraft. This in-flight program was an attempt to make a start in the direction of improved hand controls.

The controls used in this program were developed by the Hughes Aircraft Company under the sponsorship of the Air Force Flight Dynamics Laboratory. They are a result of a study which evaluated, in terms of pilot opinion, several promising controller concepts based on observation of mock-ups. This program eventually led to the fabrication of three basic configurations suitable for installation in an aircraft.

The in-flight investigation of these controllers was conducted using a Cornell Aeronautical Laboratory (CAL) B-26 variable stability airplane as a test vehicle.

The controllers were evaluated while performing various up-and-away, and landing approach, flight maneuvers. The B-26 variable stability airplane was mechanized to simulate the characteristics of three different B-1 type airplane mission configurations. The first configuration was representative of the high subsonic, medium altitude condition during which a variety of maneuvers were performed, such as obstacle avoidance turns, clearing turns, zero "g" arcs, and climbing and descending turns. The second configuration was representative of the high subsonic terrain following maneuver which provided a precision tracking task at altitude. The third configuration represents an ILS approach to touchdown and subsequent take-off.

Two US Air Force pilots and two CAL pilots evaluated each controller configuration four times, resulting in a total of sixteen evaluations per pilot and sixty-four for the complete program. Controller configurations were randomized. Each evaluation included all three B-1 type simulations and the maneuvers outlined above.

In-flight data was obtained in two basic forms for each evaluation. The first was pilot comment response to a comment guide designed to obtain pertinent information in the areas of longitudinal and lateral-directional handling qualities and cockpit environment. The pilot was

free to use as much time as he felt he needed to fully explain and give his estimate of the configuration. The second was essentially a combination of pilot ratings for the combined airplane-controller-mission as outlined by the Harper-Cooper handling quality rating scale, and oscillograph recorded controller input and airplane response data. The latter was obtained during the tracking and ILS approach tasks where deviations between command and actual aircraft motions were noted.

CONTROLLER DESCRIPTION

a. Conventional Wheel and Column: This was a B-26 design in the right seat position, and similar to the standard center-hub wheel installed in the left seat for the safety pilot. It was instrumented to provide electrical signals proportional to position and force. It was modulated by a longitudinal feel system providing various amounts of column travel per column force, breakout force and hysteresis. Mechanical springs provide lateral control force gradients. Electrical pick-offs were used to generate command signals for the variable stability system which in turn drive the electro-hydraulic actuators on the control surfaces.

b. Circumferential Wheel and Column: This test configuration differed from the standard in that the center hub was eleminated by causing a semi-circular yoke to travel through a set of upper and lower rollers contained in the top of the column. The same center of rotation was thus preserved.

c. Circumferential Wheel as above but with an Alternate Hand Controller Activated in the Form of the Right Hand Grip: The column itself was capable of being locked while the alternate controller was unlocked to free it from its fixed position and allow it to provide fore and aft motion for pitch, and left-right motion for roll control.

d. Dual Side-Arm Controllers: Here the configuration was quite different, with left and right hand grip controllers mounted as integral parts of a B-26 seat. The unit contained a force-feel system providing for variable force gradients, break-out forces and damping. The two grips were mechanically linked around the after part of the seat by means of a light-weight but smooth and positive system of torque rods and cables. A set of four rotary synchros were used as transducers in each of the two axes, pitch and roll, to provide a realistic system.

RESULTS

Results based on complete correlation of all data are not yet available. Flight testing was formally ended in January this year. In general the new concepts were well received by the subject pilots. The designs were considered to be well conceived and executed in terms of overall operating characteristics. Further work is indicated in optimizing both force and displacement factors and the control system gains.

A COMPUTER BASED LEARNING SYSTEM FOR INHIBITORY TELEOPERATOR CONTROL

Amos Freedy
Frederick Hull
John Lyman

Biotechnology Laboratory
University of California at Los Angeles

ABSTRACT

A concept of adaptive aiding for performance improvement in remote handling is described. The concept incorporates an autonomous control subsystem (ACS) that is able to supplement the operator's control function while operating in parallel with him. The behavior of the ACS is established through a process of learning, observing the operator's control function in relation to the environment and manipulator output. The computer based system establishes a decision making policy which is based on conditional probability. Initially, the output device is totally controlled by the operator, while the computer system acts as a passive observer. As the operation continues the computer system is gradually transformed to an active controller, thereby reducing the operator's function to that of an action initiator and inhibitor. A learning model based on maximum likelihood principles constitutes the mathematical basis for the system. Experiments have indicated that the theory is feasible. In addition to being a theoretical and engineering challenge, the development of a system which reduces the decision load of the operator and provides improved overall performance appears to fulfill

urgent need in the field of man-machine control where the information load imposed on the operator surpasses his channel capacity.

---

* This material is principally based on a dissertation which was submitted by Amos Freedy as partial fulfillment of the requirements for the degree of Doctor of Philosophy in Engineering at the University of California at Los Angeles.

The work was mainly supported by the U.S. Army Research Office, Durham, North Carolina, Grant Number: DA-ARO-D-31-124-G1035. Additional support was provided by the Department of Health, Education and Welfare, Social Rehabilitation Service, Washington, D.C., Project Number: RD-2904-MPO-70-C1.

INTRODUCTION

Human control of a complex machine system requires precise coordination in time and space. Man is limited as a controller by inherent psychophysiological restrictions such as bandwidth, neuromuscular fatigue, and his inability to maintain perceptual vigilance. Despite his apparent inadequacies as a controller, man is surprisingly successful in a variety of "special purpose" systems such as automobiles, airplanes, and earth-moving equipment. But as man's technology continues to advance, he is faced more and more with the need to develop control skills which lie close to the boundaries imposed by his inherent psychophysiological limitations.

Remote manipulation is an area of particular concern because of the increasing requirements to perform handling operations in environments which are hazardous to life. In cases such as lunar surface scraping, underwater retrieval, and the handling of radioactive materials, the problems of life support are often so great as to preclude the physical presence of the human operator in the task area. One solution is to place a manipulator device in the hazardous environment and to control its operation from another location. Such devices constitute a family of general purpose machines called remote manipulators or teleoperators (Corliss and Johnsen, 1968).

Although teleoperators protect the operator from the hazards of the task environment, they complicate his control problem greatly. An indication of this is provided by the degradation in speed of task performance. Taking the movement of the human arm as a benchmark, the most advanced teleoperator control systems are able to operate at speeds which at best are only one-fourth of normal (Goertz, 1963). More typically, task performance times are increased by an order of magnitude when teleoperators substitute for direct human manipulation (Corliss and Johnsen, 1968).

An adequate solution for the control problem was obtained by employing the master-slave concept used with "first generation" manipulators (Goertz, 1963). The device

controller, the master, is geometrically similar to the output device, the slave, and provides spatial correspondence between the states of the output and the input. This configuration offers a "natural" solution by exhibiting through the master the device's position and forces on a "built-in" one-to-one basis. However, non-visual sensory information is lost when the remote manipulator control site is mobile, such as in submarines and space vehicles where the master-slave concept cannot be used. In these cases, switches and joysticks are typically employed as controllers. With these types of controllers, the communication link between the operator and the output device is weakened. Anthropometric correspondence is eliminated; thus, kinesthetic feedback and a one-to-one input-output position mapping, and force and torque reflection are not easily achieved. The loss of other sensory cues, such as the proprioceptive information available with the master-slave system, makes the operator totally dependent upon the visual system. Distortions in the visual feedback process can not be compensated for if redundant or supportive information from other sense modes is missing. Nevertheless, distortions of some form are inevitable in the visual information utilized as feedback in remote manipulator control, as is obtained by indirect means such as T.V. display.

Other problems in remote manipulator control are due to transmission time delays. This becomes a problem when the need arises to operate teleoperators over extremely long distances, as in space exploration programs. Significant delays are introduced both by signal transmission time and by the need to process data once it has been received. Delays ranging up to 2.6 seconds occur in the operation of the moon scraper, for example. Although operation of manipulators under such delays is possible by the "move and wait" control strategy (Ferrell, 1965), the task time increases significantly.

One method of compensation for the control difficulties is to add to the control loop a subsystem, acting in

parallel to the human operator, which is able to assume a part of his decisions and control functions. "Aiding" of this type has been partly realized in a variety of man-machine control systems, as autopilots, predictor displays and tape programmed machine tools. A unique example related to manipulative devices was the Case Institute Arm-Aid which had certain motions programmed so that a handicapped operator could select a repertory of desired functions with a minimum input information (Wijnschenk, 1964). An example of an industrial programmed manipulator is the Unimate, developed by Unimation, Inc. (Balmer, 1967) which was controlled by a movement pattern record stored on a magnetic drum. A more recent approach is the work on "Supervisory Control" at M.I.T. (Ferrell and Sheridan, 1967). The concept integrates a computer which is capable of making some sub-decisions of its own to the control loop, while major goals are set by the "supervisor"--the operator.

The approach described here is somewhat more complex, presenting the concept of adaptive aiding. It involves the use of an automaton subsystem programmed on a digital computer which works in parallel with the human operator (see Figure 1), "learns" by observation the tasks being done, and progressively takes over more and more of the control responsibility. There are two main control loops; an external visual motor loop, which includes the operator's motor system, and an internal loop, which involves only the autonomous control subsystem. The manipulator can respond to either the operator or the autonomous subsystem. Initially, the autonomous control subsystem (ACS) acts as a passive observer trying to "understand" how the operator controls the manipulator and to develop an "awareness" of the environment of the operation. At this phase the ACS defines the relationship between the operator's responses and the external world. Upon the acquisition of some "experience," the internal loop can participate in the control of the manipulator. It will observe the manipulator states and other pertinent feedback and generate a future state. In this next phase the operator will act mainly as an action initiator, letting the internal loop control the manipulator. Finally, the

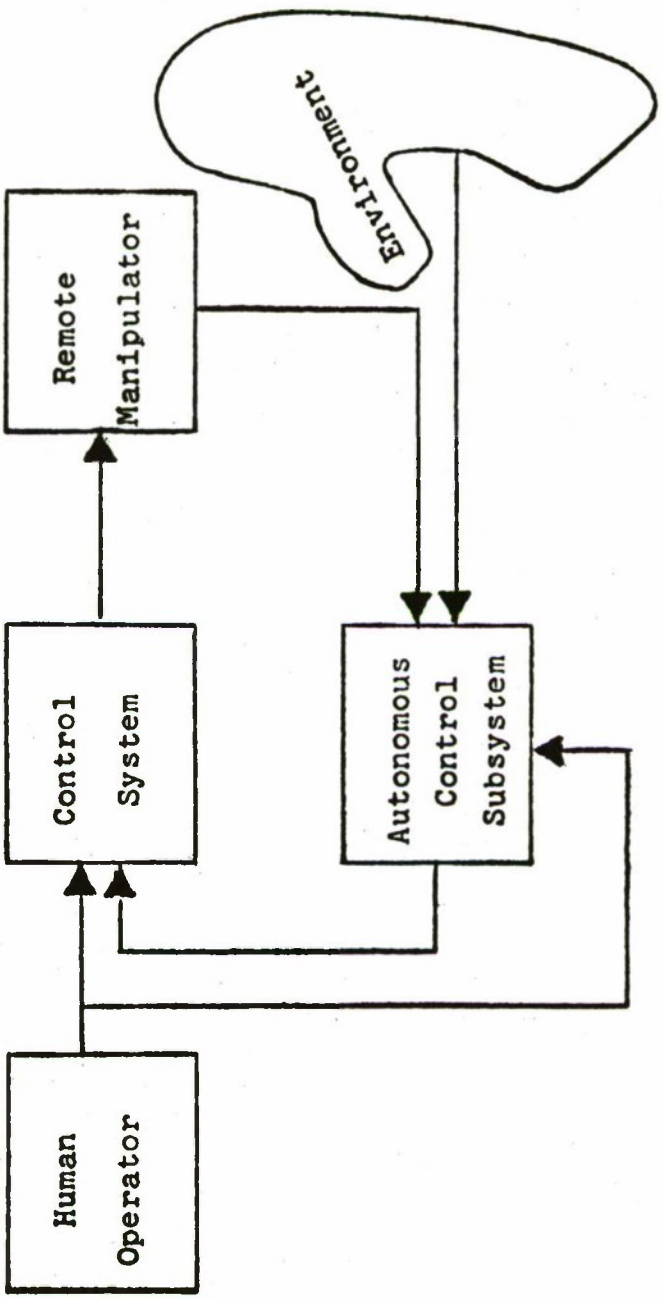


Figure 1, The Proposed Control System

function of the operator will be reduced to that of action initiator and inhibitor. Under such a configuration the decision load associated with controlling the manipulator should be substantially reduced. The ACS was designed to be capable of forgetting what it has learned and learning new tasks in place of forgotten ones. This feature provides the adaptive capability to change behavior in response to changes in the operator's control policies and in the environment.

The approach was originally described by Lyman and Freedy (1967) and was termed, "Inhibitory Control". Development work led to the use of a Maximum Likelihood Decision system as the basis for the learning automaton. This approach replaced the correlation technique suggested in the original paper; however the basic concept of control by inhibiting the actions of the automaton subsystem has been preserved.

THE AUTONOMOUS CONTROL SUBSYSTEM

The theoretical basis for the Autonomous Control Subsystem is the maximum likelihood decision principle and its structural organization is based on the Conditional Probability Computer (Uttley, 1956). It was applied to the system under the assumption that the spatial movement pattern of the manipulator is non-random. There appear to be favored paths of movement which are determined by the optimization criteria of the operator, the physical structure of the output device and the environment (Nachshon, 1965). Thus with the manipulator in a given state certain future states are more likely than others. The behavior of a manipulator can be explicitly defined by the following postulates:

1. State transitions of manipulator end point are related non-deterministically to the environment and to the manipulator's present, past, and terminal states.

2. Occupation frequencies by the manipulator of specific terminal locations within its bounded space of movement are not distributed uniformly.

The first postulate considers the phenomenon that given the environmental state $\overline{Y}_n$, present state $\overline{X}_n$, and most recent past terminal state $\overline{X}_{n-1}$ it is possible to predict with known probability which future terminal state a_{n+1} the device will approach. Given a vector $p_n(a)$ whose i'th component is the probability of occupying the i'th state in the state space, a joint probability function of transition to position $a_{(n+1)}$ will exist as:

$$P(a_{n+1} \mid \overline{X}_n, \overline{X}_{n-1}, \overline{Y}_n) \tag{1}$$

where a_{n+1} is a point in a 3-dimensional space whose components are three random variables. The second postulate considers the phenomenon that some states are more likely to be occupied by the manipulator than others. A joint a priori distribution function $P(a_{n+1})$ exists such that:

$$P(a_{n+1}) \neq \text{constant} \tag{2}$$

Following these assumptions a model for the ACS will be developed. The model can be described as a recursive relationship:

$$a_{n+1} = \underline{\Phi}(\overline{X}_n, \overline{Y}_n, P_n) \qquad (3)$$

where $\overline{X}_n$ is a vector describing the states of the remote manipulator at its n'th terminal position; $\overline{Y}_n$ is a vector describing the environmental states; and P_n is a matrix describing the state of the system, the components of which measure the experience acquired by the ACS. a_{n+1} refers to the terminal position of the output device at the $(n+1)^{'th}$ trial. The system and the input state space are discrete spaces. This limitation provides simplicity in the mathematical manipulations and in the practical simulation of the system and resembles more closely the conditional probability computer.

The constraints on the system may be explicitly described as follows:

1. The bounded space of movement of the manipulator is divided into n-discrete cells. Each cell in the space is defined as a state of the system.

2. The Autonomous Control Subsystem can only generate or respond to states which are defined by the set of n-discrete cells.

3. The environmental input and the states of the system are defined by a binary vector of m-components where each component is an all-or-none description of an input event.

In light of these constraints the following observations can be made about the functions of the ACS. The input to the ACS at the n'th trial is a binary vector $\overline{X}_n$ consisting of a set of components, $x_1, x_2 \ldots x_K$, equal to either zero or to one representing the state of the output and the environment. The environmental state consists of a map of the work space representing location of objects and obstacles. The output is a vector a_{n+1} describing the location of the cell to which the future state at (n+1) belongs.

The process of predicting the (n+1)'th state can be defined here as a mapping of a set of inputs described by a vector $\bar{X}(n)$ onto an output a_i, i=1,2...k..R, where a_k is the selected future state out of a group of R discrete states. The ACS is essentially a classification or a decision mechanism whose decision space consists of R categories from which the most probable future state a_k has to be selected.

Decision Model

An optimum decision mechanism for the Autonomous Control Subsystem was constructed using the a priori and conditional probabilities defined in postulates 1 and 2. For a decision mechanism to be optimum the average loss associated with the classification should be minimal. Define a loss function $l(a_i|a_j)$ which exists for i=1,2....R and j=1,2....R (Nillson, 1965). The loss function assigns a loss to the system when it selects a state a_i as a decision outcome, while the input vector actually belongs to state a_j. Given the conditional probability $P(a_j|\bar{X}_n)$ the conditional average loss can be written as:

$$L_x(a_i) = \sum_{j=1}^{R} L(a_i|a_j)\, P(a_j|\bar{X}_n) \qquad (4)$$

When an input pattern $\bar{X}_n$ is presented to the decision system, its decision will be optimum if the state, a_i, which is selected yields the minimum average conditional loss. Given a pattern $\bar{X}_n$, $L_x(a_i)$ is calculated for all a_i's i=1,2....R, and the optimum a_{i_o} is selected such that $L_x(a_{i_o}) \leq L_x(a_i)$ i=1,2....R.

The conditional probability in Equation (4) can be written in terms of Bayes' rule:

$$P(a_j|\bar{X}_n) = \frac{P(\bar{X}_n|a_j) \cdot P(a_j)}{P(\bar{X}_n)} \qquad (5)$$

Substituting Equation (5) into (4) the conditional average loss can be written as:

$$L_x(a_i) = \frac{1}{P(\bar{X}_n)} \sum_{j=1}^{R} L(a_i|a_j)\, P(\bar{X}_n|a_j)\, P(a_j) \qquad (6)$$

A criteria for decision is obtained by selecting a particular loss function to weigh the alternatives associated with losses in the various output categories of the process at hand. With respect to erroneous decisions associated with manipulator movement control, it would be logical to select a loss function which assigns higher losses to collision courses. Such a loss function can be expressed as:

$$L(a_i|a_j) = C_i(n)|(1-\delta_{ij}) \qquad (7)$$

where $C_i(n)$ is a parametric factor which is defined for i=1,2....R, and is a function of the position of the manipulator and the environmental state at the n'th trial. It is defined as:

$$C_i(n) = \begin{cases} 1 - \text{if no obstacle exists in the trajectory between the position at the n'th trial and the i'th position} \\ K - \text{if an obstacle exists between the present position and the i'th position} \end{cases}$$

δ_{ij} is the Kronecker delta function whose value is unity when i=j and is zero otherwise.

Substituting the loss function shown in Equation (7) into Equation (6) the conditional average loss can be written as:

$$L_x(n,a_i) = C_i(n) \cdot \left[1 - \frac{P(\bar{X}_n|a_i)P(a_i)}{P(\bar{X}_n)}\right] \qquad (8)$$

$L_x(n,a_i)$ is indicating the average conditional loss at the n'th trial. From expression (9) it is evident that $L_x(n,a_i)$ would be minimized if the system

$$\text{MAXIMIZE} \left[C_i(n) \cdot \frac{P(\bar{X}_n | a_i)\ P(a_i)}{P(\bar{X}_n)} \right] \qquad (9)$$

An optimum decision can be made by calculating the value of expression (9) for all values of a_i, i=1,2....R, and selecting the a_i for which it is maximum. Such a decision also minimizes the probability of an erroneous classification as was shown earlier. Such a decision rule is known as a Maximum Likelihood Decision.

Application of the Maximum Likelihood Decision rule to the realization of the Autonomous Control Subsystem can be done as follows: the input pattern is a vector $\bar{X}_n$ where

$$\bar{X}_n = \begin{bmatrix} x_n1 \\ \vdots \\ x_nK \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 0 \\ \vdots \\ \vdots \end{bmatrix} \qquad (10)$$

Each of the vector's K components describe a possible state of the system. A sum of 2R components is assigned to all possible terminal states of the manipulator at the n'th and (n-1)'th trials since R components are assigned to each of the R discrete cells in the state space. The environment is represented by its occupation status. (A value of 1 indicates that a particular cell is occupied.) Other components of the input vector represent the states of the claw open or closed.

The events $[x_n1.....x_nk]$ which are represented by the input pattern can be considered independent thus the

conditional probability $P(\bar{X}_n|a_i)$ can be written as:

$$P(\bar{X}_n|a_i) = \Pi p\,(x_{nj}|a_i) \cdot \Pi q(x_{nj}|a_i) \qquad (11)$$

$$j\varepsilon x_{nj} = 1 \qquad j\varepsilon x_{nj} = 0$$

where $q(x_{nj}|a_i) = 1-P(x_{nj}|a_i)$. Substituting equation (11) in the functional in (9) results in the expression:

$$\text{MAXIMIZE}\left\{\frac{C_i(n)}{P(X_n)}\left[\Pi p(x_{nj}|a_i) \cdot \Pi q(x_{nj}|a_i)\right] p_n(a_i)\right\} \qquad (12)$$

$$j\varepsilon x_{nj}=1 \qquad j\varepsilon x_{nj}=0$$

define for i=1,2....R.

The factor $\frac{1}{P(\bar{X}_n)}$ is common to all rows and therefore can be deleted.

A general expression for the decision process can be developed by introducing a RxK matrix referred to as the P matrix.

$$P_n = \begin{bmatrix} P_n11 & P_n12 & P_n13 \cdots\cdots\cdots & P_n1K \\ P_n21 \cdots\cdots\cdots\cdots\cdots & & & \vdots \\ \vdots & & & \\ \vdots & & & \\ \vdots & & & \cdots P_nRK \end{bmatrix} \qquad (13)$$

The P_n matrix is a stochastic nonstationary matrix whose elements are the conditional probability $P(x_nj \mid a_i)$ at the n'th trial where x_nj is the j'th component of input vector a_i is the i'th cell in the state space. Using the P_n matrix, the decision process can be described as:

$$I_{n+1} = \text{INDEX} \left\{ \underset{M_{a_i}}{\text{MAX.}} \left[M = [m_1, m_2, \ldots . m_R] \right] \right\} = \tag{14}$$

$$= \left\{ C_i(n) \left[[\log P_n] \cdot [\bar{X}] \left[[\log (1-P_n)] \cdot [1-\bar{X}] + [\log P_n(a)] \right] \right] \right\}$$

where the instruction "INDEX $\left\{ \underset{M_{a_i}}{\text{MAX.}} \right[M$" implies selection of the index of the maximum component of the vector. If a_{i_o} was selected as the most probable output then $M_{a_{i_o}} \geq M_{a_i}$ for all i=1,2....R.

The decision mechanism calculates the above expression and simply selects the i'th components of the output whose value is maximum. The input vectors are fed into a set of "weighing amplifiers" representing the operation performed by expression (14). The output of each amplifier corresponds to a vector component M_{a_i}. The sums are fed to the maximum selector through a threshold amplifier which compares the outputs and selects the index of the component having the highest amplitude. The threshold amplifiers are a part of the confidence test. In the confidence test the absolute "quality" of the decision is checked. This process is accomplished by comparing the value of $M_{a_{i_o}}$ to a preset threshold level θ. If the value of $M_{a_{i_o}} > \theta$ then the decision is accepted. (See Figure 2.)

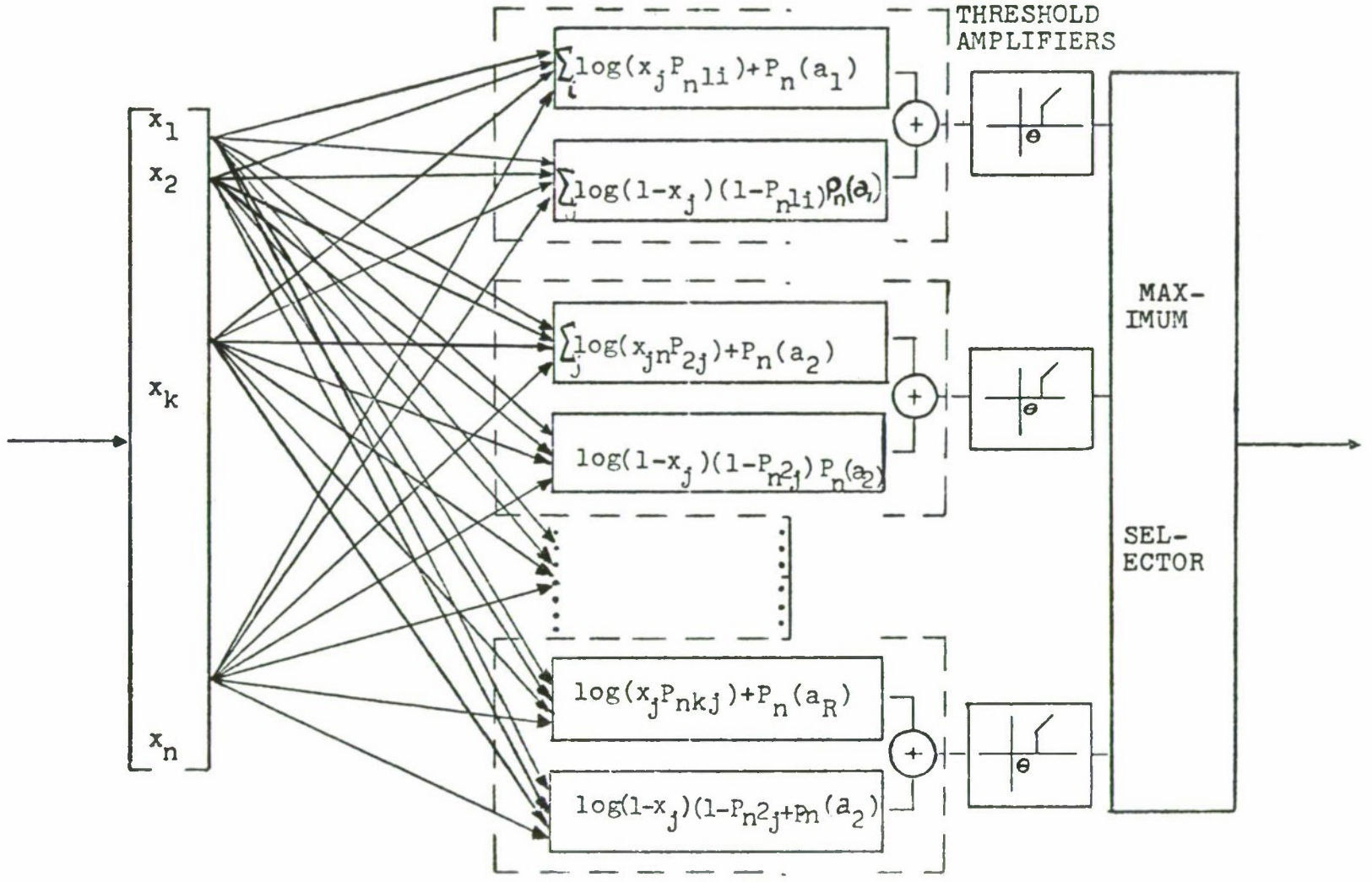


Figure 2 . The Decision Mechanism

System Training

The values of the elements of the P_n matrix and the a priori probability vector $P_n(a)$ (Equation 14) are determined by the training and experience that the decision mechanism has had in observing and controlling the manipulator. As was shown in Equation (14) these values constitute state variables influencing decision strategies of the ACS. Determination of the conditional probabilities $P_n ij$ and $P_n(a)$ is accomplished by a continuous goal directed training process. Training can be divided into two modes. The first involves passive observation of the output states and the establishment of the relationship between future states and the ACS input. At this mode training is in the form of conditioning (Uttley, 1956). The second mode of training accompanies active operation of the ACS and is applied as a response to the level of performance exhibited by the ACS. This mode involves positive reinforcement and negative reinforcement and is mainly directed toward the reinforcement of favored established patterns of behavior and the extinction of unfavored ones. Through negative reinforcement, the behavior of the system changes when environmental conditions and/or operator control strategies have changed. Negative reinforcement occurs when the probability attached to the occurrence of a particular input has changed with respect to the occurrence of a state a_i. In essence negative reinforcement causes the forgetfulness of established behavioral patterns, and is applied only when the system has actually executed its decision and the decision is rejected. As will be shown later, it might occur that a certain decision has been made but is not executed for lack of confidence.

The first mode of training is accomplished as follows: By Laplace's rule of succession (Parzen, 1960) each element of the P_n matrix or the conditional probability $P_n ij$ can be estimated as:

$$P_n ij = \frac{1 + \sum_{q=1}^{q=n} \phi(a_i \, x_n qi)}{1 + N(a_i)} \qquad (15)$$

$$\phi(a_i | x_n qi) = \begin{cases} 1 & \text{if the event } a_i \text{ occurred following the event that the input component } x_n qi=1 \\ 0 & \text{otherwise} \end{cases}$$

$N(a_i)$ presents the number of times a_i has occurred. (In terms of the manipulator system, $N(a_i)$ represents the number of times the manipulator occupied cell a_i in the total number of trials it performed.
The a priori probability component, $P_n(a_i)$ of the vector $P_n(a)$ is estimated in a similar manner as:

$$P_n(a_i) = \frac{1 + N(ai)}{2 + \sum_{i=1}^{R} N(a_i)} \qquad (16)$$

where the sum of $N(a_i)$, i=1,2....R, yields the total number of trials the system has experienced. Suppose that in the training phase an input vector $\bar{X}_n$ is applied to the ACS, and suppose that following it the manipulator is directed to state a_i. At the completion of the task the row of the P matrix which corresponds to the a_i'th state would be "updated", i.e., the conditional probability values $P_n ij$ for j=1,2....K would be changed to a new value determined by Equation (15). The a priori probability $P_n(a_i)$ will be updated according to Equation (16). In the second mode of training, positive reinforcement and negative reinforcement are both included. Training is accomplished by rewriting the expression for the conditional probability as:

$$P_n ij = \frac{1 + \sum_{q=1}^{q=n} \phi(a_i | x_n qi) \cdot \theta(a_i, q)}{2 + N(a_i)} \qquad (17)$$

where $\theta(a_i, q)$ is a function defined as:

$$\theta(a_i, q) = \begin{cases} +1 & \text{if the event } a_i \text{ occurs at the } q\text{'th trial} \\ -1 & \text{otherwise} \end{cases}$$

According to Equation (17) $P_n ij$ could theoretically approach negative values; however in practice this could not happen, since expression (17) is only used when the ACS exhibited confidence for occupying a_{i_o}.

SYSTEM IMPLEMENTATION AND EXPERIMENTAL EVALUATION

The ACS was implemented on the IBM 1800 computer and integrated to the manipulator control loop as shown in Figure 3. The operator is linked to the manipulator through a control logic stage. The function of the control logic is to interpret operator responses and decide whether to give control of the manipulator to the computer or to the operator. The manipulator is controlled by either a position mode or a rate control mode. When the manipulator is driven by the operator it is set in the rate mode, and when driven by computer it is set in the position mode. The three degrees of freedom of the manipulator are shoulder rotation, shoulder elevation and flexion-extension movement. The manipulator claw is controlled and powered by a foot pedal. The operator's control console consists of a joy stick combined with a GO/override switch and a light indicator. The joy stick switch is used by the operator to generate a "Go" signal that is to shift the control of the manipulator from the joy stick to the computer. The switch is also used to override computer position control and transfer control back to the joy stick at the completion of the computer generated movement. The light indicator at the operator's control console is used to signal the operator as to whether or not the ACS (or the computer) has the required level of confidence for executing its decision.

The operation of the system, in point control can be summarized as follows. Suppose the manipulator is approaching point a_1 in space and the operator desires to move next to a_2. As the output device reached a terminal state at a_1 the

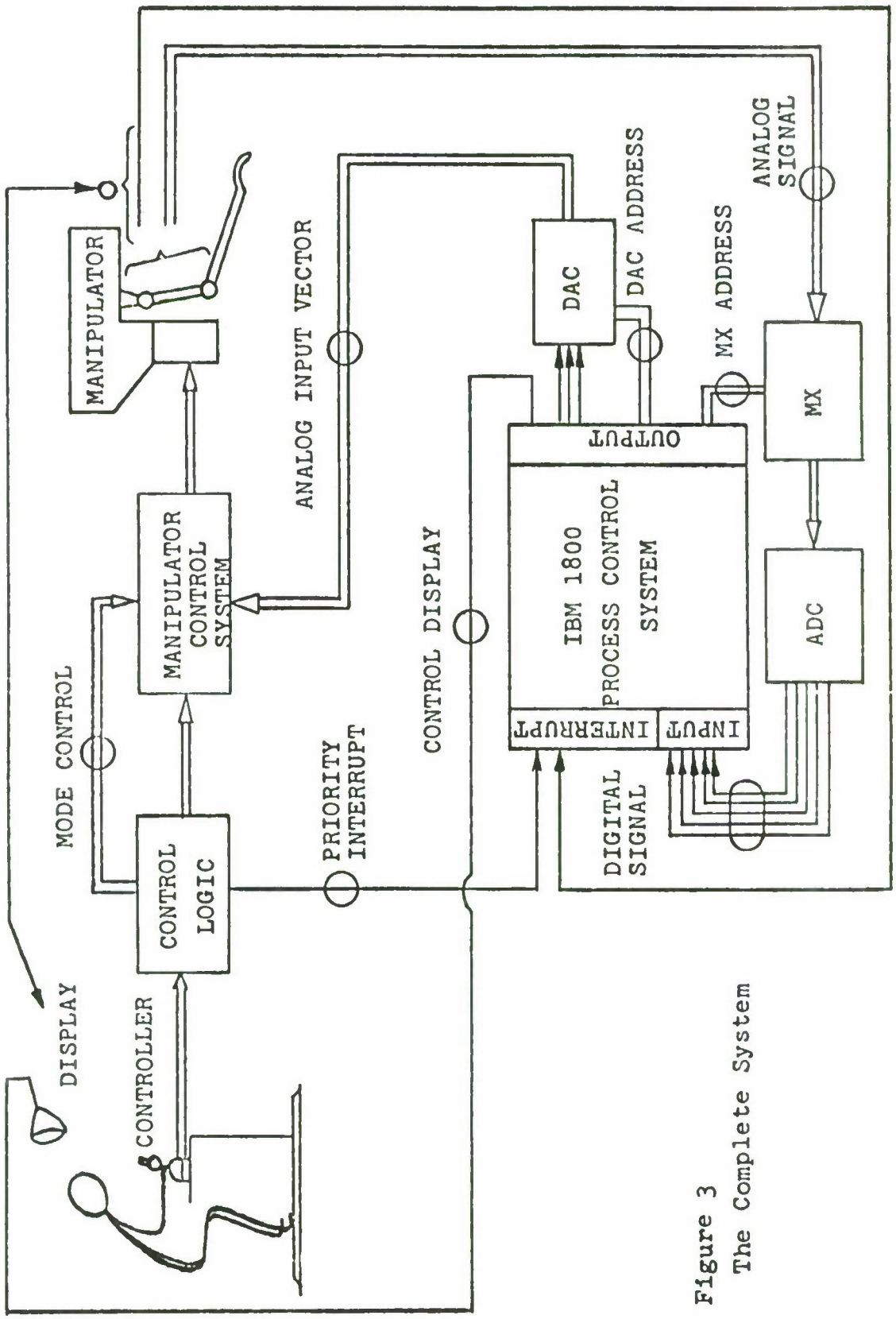


Figure 3
The Complete System

computer (of the ACS) would record the position of the arm and the claw, scan the environment, construct an output vector, and apply the decision process. If the decision outcome is accepted it would generate the selected future position and store it in a stand-by mode. Following a "GO" signal from the operator the control signals are transferred from the stand-by buffer storage and applied to the position control loops. The manipulator is directed to the most probable position selected by the ACS. An incorrect decision by the system is corrected by the operator and computer control will be cut off. The operator will take over control from the computer and direct the manipulator to the desired location. As soon as the task is completed a task completion signal will call the computer to start P_n matrix update process.

An experimental study was performed to evaluate the realized system.* The objectives of the study were to test the system's learning capabilities, and to substantiate the capabilities of the ACS to reduce the decision load of the human operator.

The study was divided into two sets of experiments. In the first set only machine performance was investigated. In the second set man-machine performance was studied. As a basis for the studies, a set of three "meaningful" manipulative tasks was selected. The question as to what constitutes a meaningful task was of major consideration in the interpretation of the experimental results. The significance of a particular task is determined by the degree to which it resembles standard manipulative tasks. With respect to the evaluation of the ACS a set of tasks are meaningful if they intersect each other, that is if they contain joint terminal states. If a number of tasks intersect each other, then the corresponding input vectors to the ACS are joint sets. This implies that each component of the input vector could infer a variety of movement patterns at different times. This property is of extreme importance in testing the system since the system must undergo learning, forgetfulness, and relearning throughout the performance of the set of tasks, and would put to test basic behavioral properties.

---

* Due to limitation of computer memory, the realized system did not include environmental mapping.

A set of three intersecting tasks were chosen. The tasks are classified as (1) Movement of objects from a random location to a fixed point, as shown in Figure 4a, (2) Movement of objects from a fixed location to a bounded area, unloading the objects to an ordered formation shown in Figure 4b, (3) A sequential movement of the manipulator between four locations--loading an object at the third location and unloading it at the fourth. The task was continued by returning back to the first location as shown in Figure 4c. The objects were all identical in size and color and consisted of 1-1/2" by 1-1/2" red cubes. Tasks 1 and 2 consisted of the relocation of 16 objects while Task 3 consisted of the relocation of 10 objects from the first location to the second and 10 objects from the third to the fourth location.

Method

Four subjects were used to test and study man-machine performance. Tasks 1, 2 and 3 (see Figure 4) served as the performance objectives. Two of the subjects performed the tasks without the aid of the computer while the other two were using the computer. The tasks were organized into blocks which are termed RUNS, each of which consists of three tasks. There were 32 trials of Tasks 1 and 2, while Task 3 consisted of 40 trials. The tasks in each RUN were performed in the sequential order of Task 1 followed by Task 2 which in turn was followed by Task 3. The subjects were instructed to perform three RUNS with two 15 minute breaks between each RUN and a 5 minute break between tasks.

All subjects in the group were highly motivated and were previously exposed to the development of the system; however, none had any experience in operating the system or

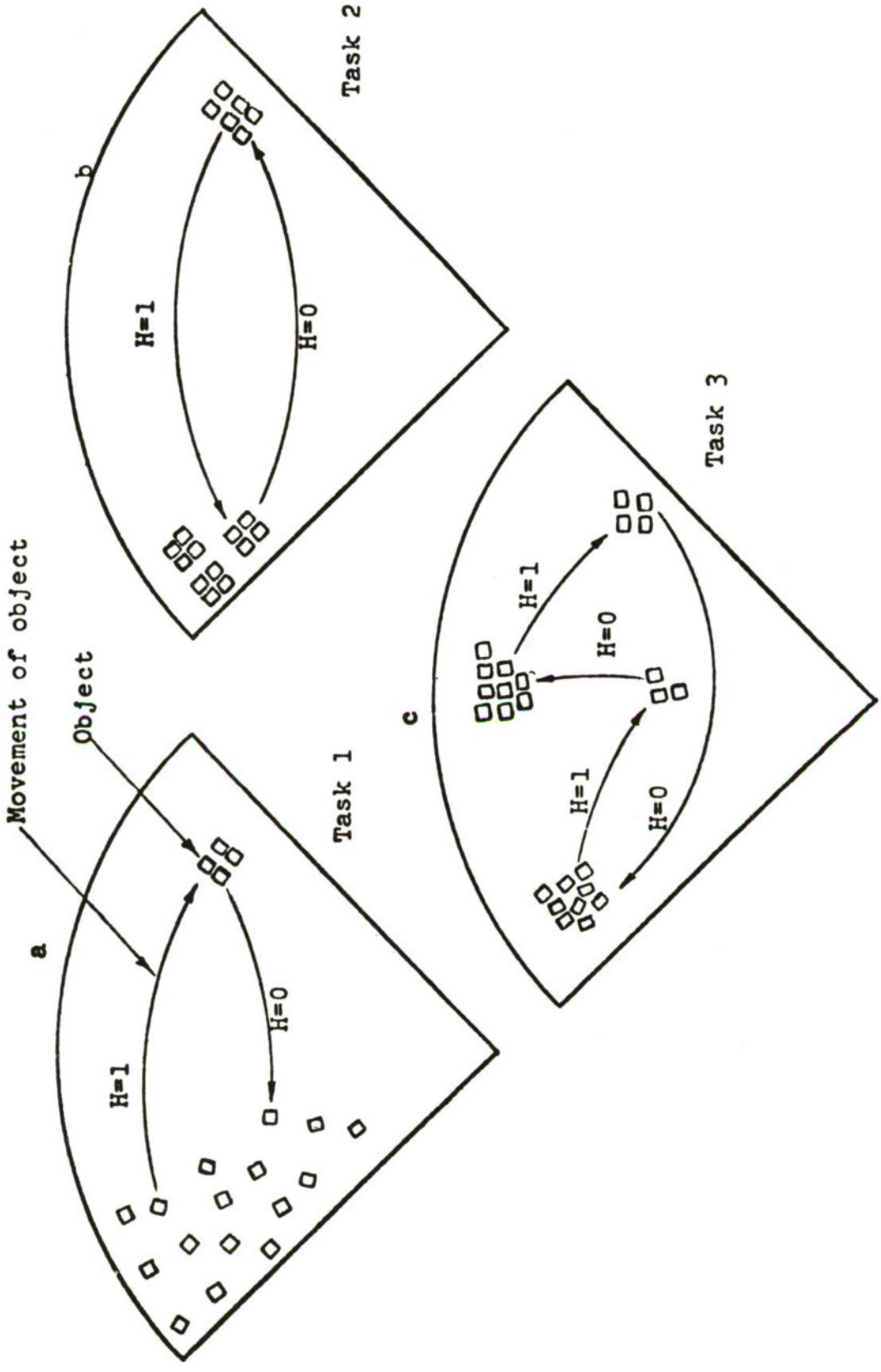


FIGURE 4 ,The Experimental Tasks

knew any details of its operational procedure. Each subject was given a 45 minute training session followed by a 15 minute break before beginning RUN 1.

The computer group was instructed to use the computer after the first 10 minutes of manual training; however, total training time for the two groups was identical.

The tasks performed in the training session were randomly selected and consisted mainly of positioning the manipulator at random locations, and picking up objects at will. Measures of performance were taken both for the man and for the machine. Trial time completion was used as a basic measure of performance. Trial time is defined as the time segment between object pickup and object release. Trial times were recorded using the claw's "cat whisker" as an indicator for trial initiation and trial completion.

In the computer group, trial time was divided into two segments, (1) computer control time and (2) override time. The computer time is the portion of overall trial time in which the computer controlled the manipulator. Override time is the portion of time that the manipulator was under operator control. Total trial time was equal to the sum of the computer time and override time.

Two measures can be derived from the data in order to evaluate system performance. First is a comparison between the time performance curves of the two groups. That is, inspecting whether or not the computer group has any advantage in total task completion time. The second measure describes the level of information "unburdening" of the human operator by the computer. That is the amount of time at each trial that the computer controls the output device. Had the computer not been in the loop, this active control time would have required the human operator.

Results

Direct comparison between the performance of the computer group and the control group in mean trial time as a function of trials is represented by the "learning curves" in

Figure 5a, b, c. Each of the curves corresponds to a task. The ordinate of each curve represents a sliding mean calculated in steps of 10 trials each over the group of 2 subjects. The abscissa simply represents the number of trials; however, the RUN number is also indicated. Curve (a) corresponds to Task 1, curve (b) to Task 2 and (c) to Task 3. The dashed line of each curve represents the sliding mean for the control group while the solid line represents the mean of the computer group.

A sample measure of the system's capability to unburden the information control load of the operator can be obtained from Figure 6. The figure shows mean task time (for Task 3) of the actual period in which the operator was actively controlling the manipulator. Graph A represents the mean trial time of the group of 2 subjects as a function of trials. The upper curve (which is guided by the circles) represents the mean trial time of the control group. In essence it represents the overall time that was required by the operator to control the manipulator in the performance of each trial. The lower curve represents the mean time of active control, or crude time required by the subjects to perform a trial. This time is termed override time and describes the time in which the operator is assisted by the computer in performing the task. Graph B represents the mean override time calculated over all the trials of Tasks 3 for the computer group. Graph C shows the mean active control trial time for the control group or the group without the computer. The shaded part of Graph C superimposes the mean override time on the mean trial time of the control group.

The rates at which control over the manipulator is transferred from the operator to the computer is shown in Figure 7. The ordinate represents the percent of mean time the computer or the operator controlled the manipulator. The abscissa represents the number of trials. In Graph A the curve which is guided by the circles represents the percent of override time or active control time averaged over the computer group in performing Task 3. The curve which is guided by the triangles represents the percent of computer control time. Graph B illustrates in its amplitude the mean percent of override control time at the last 10 trials. The

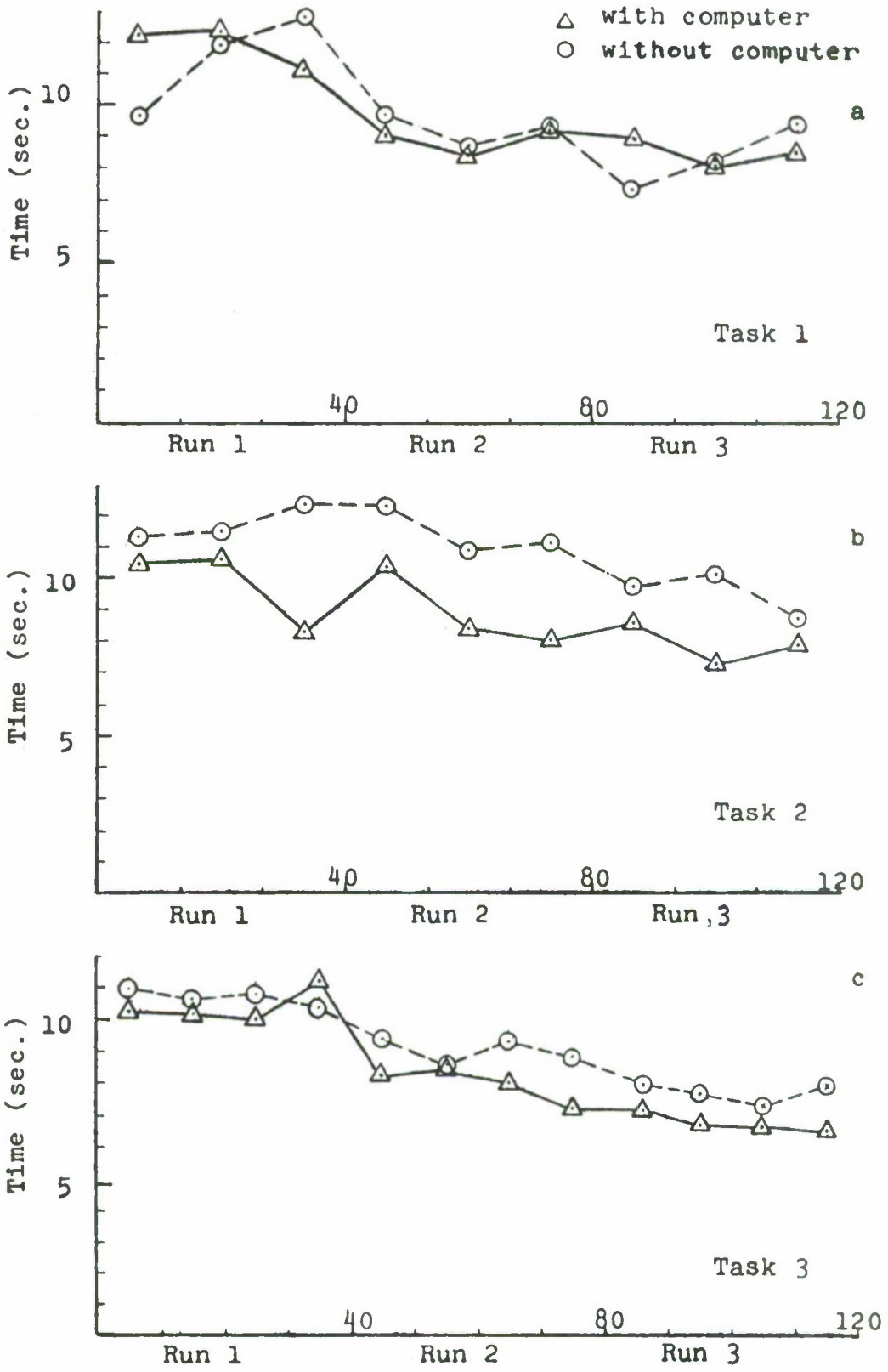


FIGURE 5. Operators Learning Curves

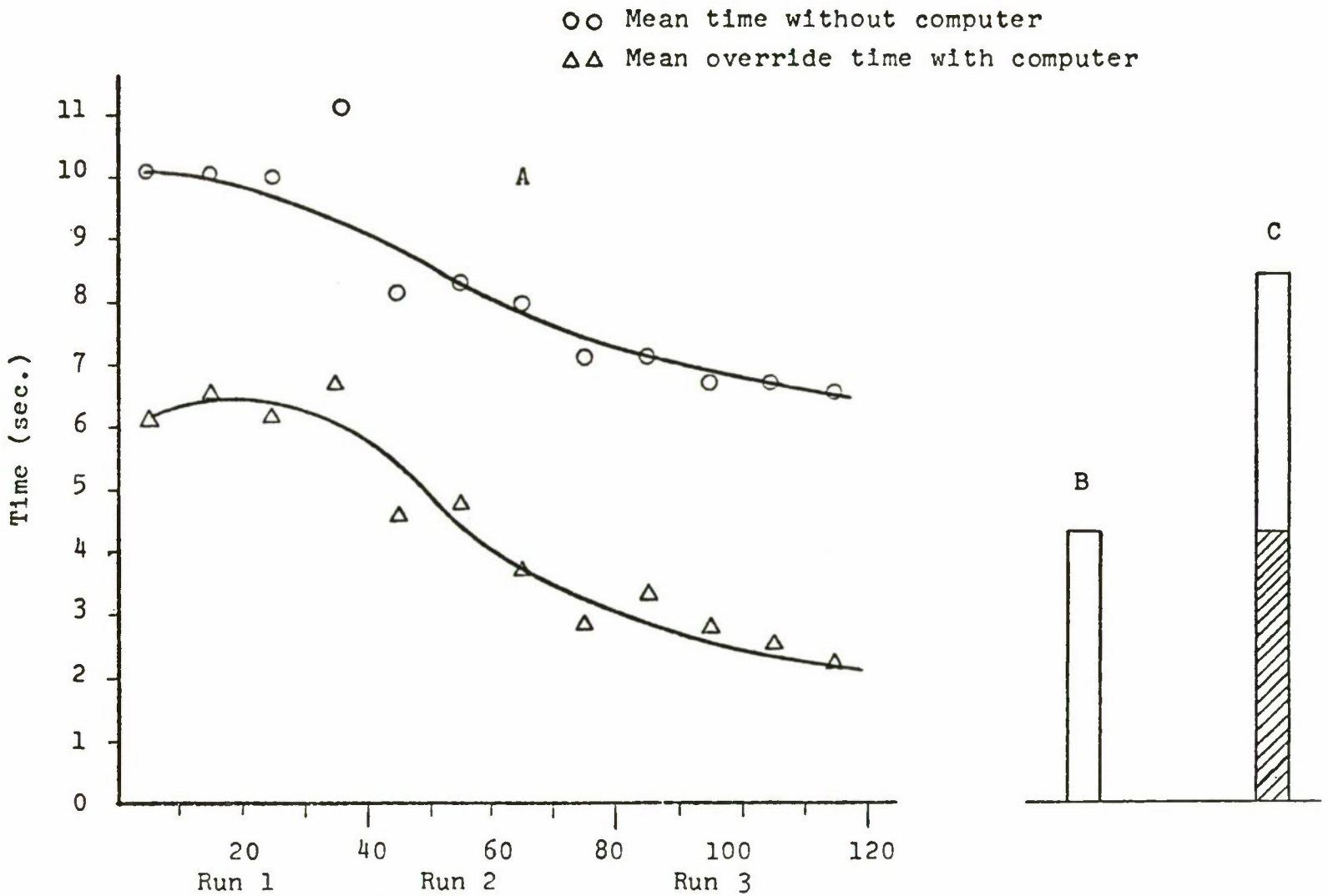


FIGURE 6 Mean Time of Operator Control

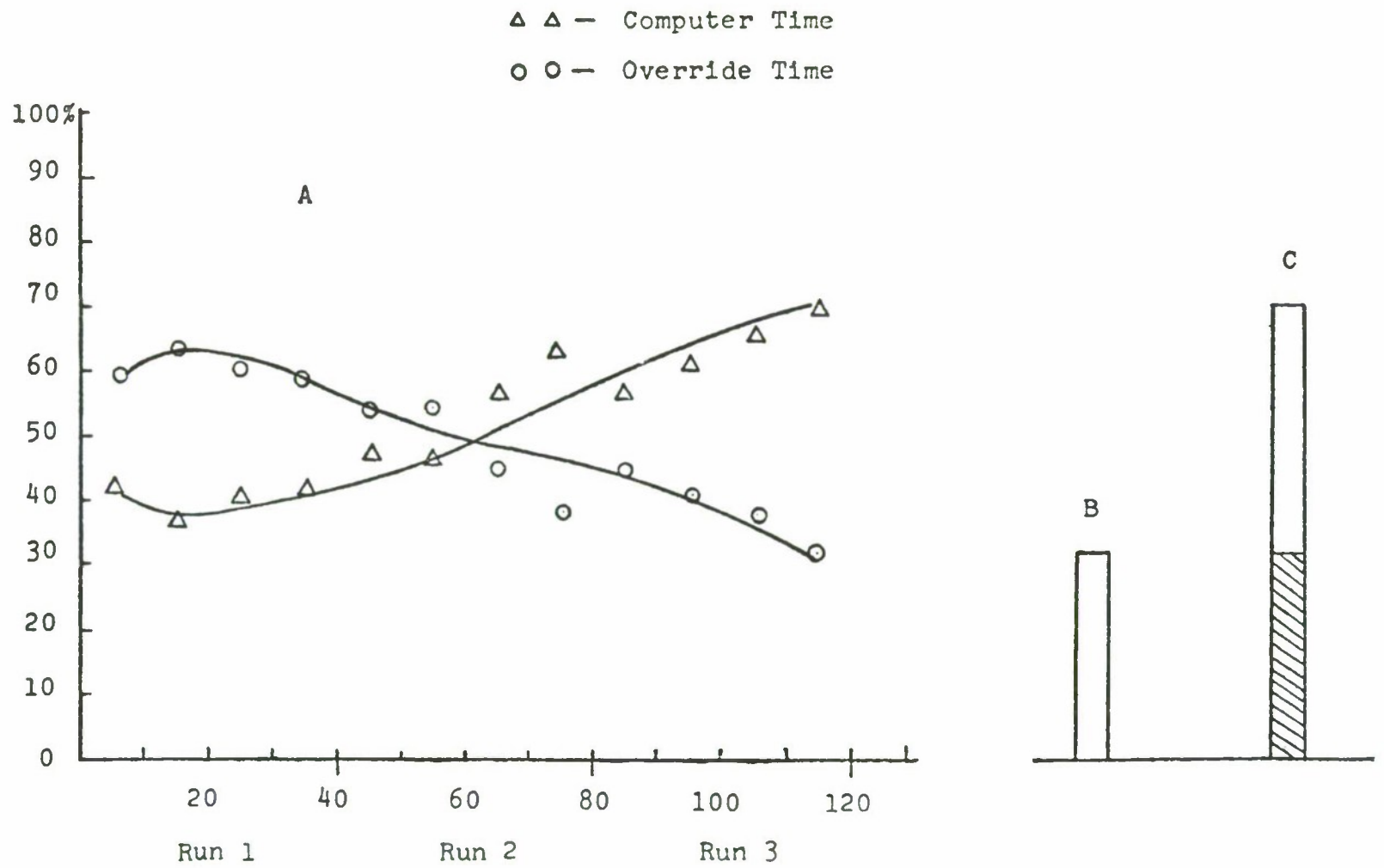


FIGURE 7 Distribution of Trial Time Between Operator & Computer

amplitude of Graph C describes the mean percent of computer time for the last 10 trials.

Discussion

Generally the computer aided system offered an advantage in achieving lower performance time in Tasks 2 and 3 but no advantage in the performance of Task 1. This result is not surprising since Task 1 is based on random movements (as was shown in Figure 4). The redundancy in the tasks is too little to be captured by the system. Thus the task's trials could hardly become a decision outcome. It is worthwhile to note that at the end of RUN 3 the computer did capture some movement pattern and distinctly offered some assistance.

In the performance of Tasks 2 and 3 the computer offered distinct advantage which increased with the number of trials. It is interesting to note that in both tasks the mean trial time was shorter at the beginning of RUN 1. This advantage can be attributed to the intersection between the tasks as discussed earlier. An important factor in the interpretation of the results is the inherent decision time delay of two seconds which was partly added to the mean trial performance time of the computer group. Had the decision been faster, undoubtedly the overall trial performance time would have been shortened also. This conclusion can be drawn from the observation of the actual operation of the system by the computer group. As the operators' skill improved they actually waited for the decision to be made.

The subjective evaluation by the computer group of the level of aiding that it received was positive. The first subject indicated that the ACS significantly helped his performance in Task 3 which was the most complex. In the simpler tasks he found little advantage in using the system. The evaluation of the system by the second operator was more positive. He claimed to receive considerable assistance from the system in all three tasks.

Figures 6 and 7 clearly illustrate the capabilities of the system for unburdening operator decision load. Since the information load of the operator is higher when he is

actively controlling the manipulator than when he is simply observing the computer's performance. Any reduction in active control time reduces his information load.

Graphs 6B and 6C indicate that the mean trial time of active control was cut for the computer group by 50%. This amounts to a significant reduction in operator decision load. Direct conversion between reduction of active control time to reduction in information load is beyond the scope of this work; however, intuitively one could expect a correlated reduction. Figure 7 illustrates the capability of the computer to actually release the operator from active control as the system gains experience.

Initially the computer controls only 42% of each trial while the operator controls 58%. At the middle of RUN 2 the control is equally shared between the operator and the computer. At the end of RUN 2 the computer starts to take over a large share of the work load approaching a level of 68% while leaving the operator with only 32% of the work load. In essence, as the machine learns to perform a task the operator's function approaches that of an action initiator and supervisor.

SUMMARY

A theoretical learning model was used as a basis for the development of an adaptive aiding system for the human operator in a control loop. Central to the aiding system was the concept of the Autonomous Control Subsystem (ACS), a subsystem connected in parallel to the control loop and capable of making part of the required decisions for controlling the output device. The ACS was able to partly unburden the operator's decision (or control information) load or supplement for his inherent limitations as a controller.

Due to technological and economical limitations the scope of the realized system was limited compared to the theoretical model; however, its performance indicated the validity and feasibility of the concepts involved. Experimental evaluation studies of the realized system revealed both its capability to improve overall performance and to significantly reduce the operator's decision load.

The present system introduces a new dimension to the man-machine relationship--cooperation of the operator with a machine component which autonomously generates behavior. This aspect of the concept must be further explored before optimum allocation of control function can be made. In addition, practical control situations will most likely involve remote viewing, and may also involve time-delays, such as those encountered in extraterrestial transmission. The influence of these factors will have to be examined. The objective will be to develop controls, displays, and an operating strategy which allows the operator to take maximum advantage of the help promised by adaptive aiding in various applications.

BIBLIOGRAPHY

Balmer, T. R., "Unimate," Design News, 22:36-38, July 5, 1967.

Corliss, W. R. and E. G. M. Johnsen, Teleoperator Controls, NASA Office of Technology Utilization, NASA SP-5070, Washington, D.C., December 1968.

Ferrell, W. R., "Remote Manipulation with Transmission Delay," IEEE Transactions Human Factors in Electronics, HFE-6(1), September 1965, pp. 24-32.

Ferrell, W. R. and T. B. Sheridan, "Supervisory Control of Remote Manipulators," IEEE Spectrum, Volume 4, Number 10, October 1967.

Lyman, J. and A. Freedy, "Inhibitory Control; Concept of a First Model," Third Annual NASA - University of Southern California Conference on Manual Control, NASA Report SP-144: 311-314, 1967.

Goertz, R. C., "Manipulators Used for Handling Radioactive Material," Human Factors in Technology, McGraw-Hill, New York, 1963.

Nachschon, A., Skill Acquisition in Three Dimension End-Point Control, M. S. Thesis, University of California, Los Angeles, July, 1965.

Nilsson, N. J., Learning Machines, McGraw-Hill, New York, 1965.

Parzen, E., Modern Probability Theory and its Applications, John Wiley and Sons

Uttley, A. M., "A Theory of the Mechanism of Learning Based on the Computation of Conditional Probabilities," International Congress of Cybernetics, Gauther Villar Press Namur, 1956.

Wijnschenk, M. J., "Engineering Evaluation and Preliminary Studies of the Case Research Arm Aid," Report Number EDC 4-64-3, Medical Engineering Laboratory, Case Institute of Technology, Cleveland, Ohio, 1964.

ON THE POSSIBILITIES OF TACTILE INFORMATION TRANSMISSION FOR THE USE IN ARM PROSTHESES*

H.G. Stassen A.W.A. Meyer A. van Lunteren

Man-Machine Systems Group of the Laboratory for Measurement and Control, Department of Mechanical Engineering, Delft University of Technology, The Netherlands.

0. Abstract

In using prostheses for upper extremities, an artificial information channel to transmit information, for instance of the grasping function of the hand to the patient can be created by tactile stimulation of the skin. Due to weight, power expenditure and size a device with only one stimulation unit has been used. A frequency or amplitude modulation system has been applied.

To start with, the just-perceptible threshold values and the just-noticeable differences determined are reported.

To understand in what way an optimal information exchange between the prosthesis and the wearer can be obtained a series of tracking experiments have been carried out. In particular, the interest was focussed on (i) the influence of non-linear operations on the input of the tactile display, as well as on (ii) the problem of presenting a reference point.

It is shown that good information exchange is possible within a frequency range of 5-80 Hz and an amplitude range of 0.1 - 1.0 mm. Furthermore, it is shown that in tracking experiments a three-state presentation is very useful. The use of a dead zone about the reference point enables the prosthesis wearer to minimize the drift with regard to the reference point.

A comparison between the equivalent information presentation on the one hand by tactile displays, and on the other hand, by visual or auditory displays has been carried out.

1. Introduction

In using prostheses for the upper extremities a mutual exchange of information between the wearer and his prosthesis is necessary for precise functioning. Hence, to control the prosthesis correctly, the wearer needs to be informed about

*To be presented at the 6th Annual Conference on Manual Control, april 7-9, 1970, Wright-Patterson AFB, Ohio.

the state of his artificial arm. This is accomplished by the feedback of the external forces acting on the stump, which normally results in a rough control of the prosthesis. However, to obtain information about the internal forces, for instance the grasping function of the hand, an artificial information channel to the wearer should be created.

Because the handicapped person is in general constrained to head labor, the eyes and ears should not be used as information transmitters. Therefore, a study has been made to answer the question, whether the skin can be used for information transmission. Because thermal stimulation cannot be used in the frequency range required for prosthesis control, only the tactile stimulation of the skin has been investigated. A detailed description can be found in the Annual Report of the Man-Machine Systems Group [1].

From the many publications, reporting experiments of tactile stimulation, the following parameters of interest can be deduced (1) frequency; (2) amplitude and (3) shape and size of the stimulator.

With regard to the first two parameters, i.e. frequency and amplitude, some data have been reported by Wilska [2]. He reported just-perceptible amplitudes for stimulation with a 50 Hz sinusoidal signal at different parts of the body, for instance: finger tips 2.0 μm; posterial side of upper arm 21.0 μm. Other investigations with reference to the just-perceptible stimulus were executed by Hugony [3] and Verillo [4]. Unfortunately, there is a rather big variety in the upper threshold values, i.e. the pain threshold. Reported data are found by Hugony, Von Békésy [5], and Keidel [6], who found respectively 50 dB, 60 dB and 80 to 100 dB for the ratio between upper and lower threshold values. Starting from the data of Wilska [2] for the just-perceptible stimuli and from the pain threshold values just-mentioned, the maximum amplitude which is tolerable, should be about 5 mm, a value which is much too high for continuous stimulation in prosthesis control. Just perceptible threshold values as function of the frequency are reported by Eijkman and Vendrick [7]. They found that most of the receptors sensitive to touch have dynamic characteristics, however, the bandwidth is limited for the lower frequencies as well as for the higher ones. Eijkman found that the time of adaptation to a certain stimulus for a group of receptors is approximately 30 msec resulting in a lower limit of 5 Hz, while the maximum frequency for a single fibre unit, due to the recovery time, is about 250 Hz. Only for high pressure resulting in skin displacements of 1.5 mm or more a constant sensation is achieved. However, this may result in tissue damage due to poor blood circulation, hence, these large skin displacements cannot be applied in prosthesis control. According to Goff [8], the ability for discrimination is rather poor above 70 Hz; a result which later on was affirmed by Anderson, Attell and Munson [9].

Interesting experiments have been executed by Verillo [10] with reference to the size of the probe. He found that for surfaces < 0.02cm^2 the just-perceptible threshold values are independent of the frequency, while for surfaces < 0.32cm^2 the threshold values are only independent for frequencies < 70 Hz.

No publications could be found which indicate whether this is also valid for stimuli higher than the threshold value. Verillo [10] reported that the shape of the probe has no influence on frequency and amplitude sensation. Finally it should be noted that Eijkman and Vendrick [7] show that for frequencies above 15 Hz dependent on the amplitude sometimes the skin is not able to follow the probe, while Verillo [10] affirmed a loss of contact between probe and skin above 25 Hz.

In literature, two assumptions are given to describe the relation between stimulus and sensation, viz. the law of Weber-Fechner [11] and Steven's law [12]. The law of Weber-Fechner can be denoted as:

$$S = k_1 \ln \frac{I}{I_o}, \tag{1}$$

and Steven's law as:

$$S = k_2 (I - I_o)^q. \tag{2}$$

Here the quantities k_1, k_2 and q are constants, the quantity S is a measure for the sensation, and finally the quantities I and I_o are the intensity of the stimulus and the just-perceptible threshold value respectively. For tactile stimulation, Stevens found the following numerical values: amplitude sensation for a frequency ν = 60 Hz results in q = 0.95 and for ν = 250 Hz in q = 0.60.

From the literature reported before it follows that information exchange might be possible with sinusoidal stimulation within an amplitude range of 0 to 1.0 mm, and within a frequency range of 5 to 80 Hz. Furthermore, for a mechanical stimulator it will be preferable if the surface of the probe is smaller than 0.32cm^2, while the shape of the probe is not important.

2. Threshold values for tactile stimulation

The tactile stimulation of the skin has been investigated using two different types of stimulators, viz. an airchamber fixed to the skin and a mechanical stimulator. The first type can be considered as a variant of the airjet system of Bliss and Linvill [13], where the airjets are replaced by

an airchamber, in such a way that the skin can be used as a membrane. Frequency modulation between 5 and 100 Hz rather than amplitude modulation promises good results. The second type of stimulator is an electrodynamical position controlled vibrator, so that the influence of the mechanical skin impedance on the dynamics of the vibrator could be neglected. This vibrator can be applied in the frequency range of 5 to 80 Hz and in the amplitude range of 0 to 1000 µm. Its probe has a spherical shape, so that the stimulator can be used for a perpendicular as well as for a tangential stimulation of the skin. The probe diameter is 2mm, while the fitting consists of a ring with an inside diameter of 25mm, in order to achieve that a defined skin area can oscillate, similar to what Eijkman and Vendrick [7], and Verillo [4] have done.

By using the airchamber the threshold values for amplitude have been determined. Therefore, the amplitude of a sinusoidal test signal for a series of frequencies has been increased or decreased from below or above the threshold value. The stimulator was placed on the volar side of the lower arm. The experiments were executed with three male subjects. As an example the results obtained from one subject are given in Fig. 1. In all measurements the values

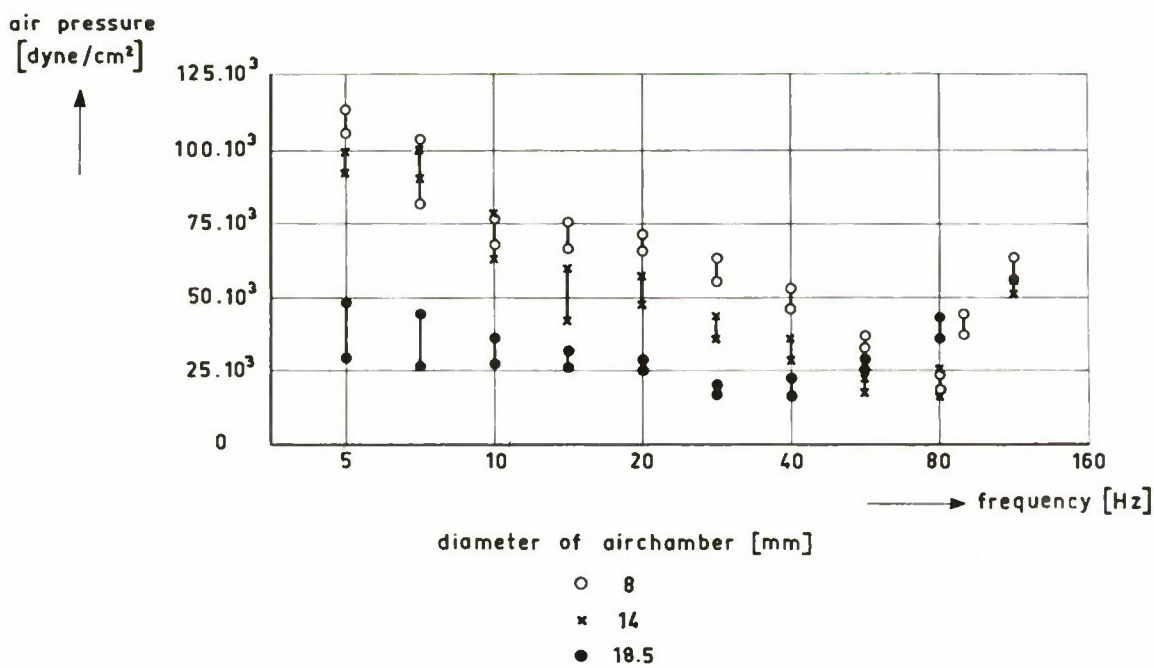


Figure 1: Threshold values of the skin for pneumatic stimulation as a function of frequency for three values of the diameter of the airchamber.

found by increasing the amplitude are higher than those found by decreasing the amplitude. The figure shows also the influence of the probe diameter. In particular, it is shown that the frequency belonging to the lowest threshold value decreases when the diameter increases. An explanation for this effect can be found by studying the mechanical impedance of the skin. Since in the airchamber a pulsatory pressure will be built up, the skin will start an oscillation. Define now the impedance z of the skin as:

$$z = \frac{p}{v}, \tag{3}$$

where the quantity p means the amplitude of the pressure in the chamber and v means the velocity amplitude of the skin. Based on results of Oestreicher [14] concerning the impedance of an oscillating sphere in a visco-elastic medium, Von Gierke et al [15] derived a simple expression for the impedance of the skin in the low frequency range:

$$z = 6\{(\sqrt{\rho\mu_1} + \frac{\mu_2}{a}) + j(\frac{1}{9}\rho\omega a - \frac{\mu_1}{\omega a})\} \tag{4}$$

where ω = the frequency in rad/sec;
a = the radius of the stimulator in cm,
ρ = the specific mass density in g/cm^3,

and where μ_1 and μ_2 are constants. Von Gierke has found that $\rho = 1.1$ g/cm^3; $\mu_1 = 2.5\ 10^4$ dyne/cm^2 and $\mu_2 = 150$ dyne sec/cm^2. In Fig. 2 the impedances of the skin are shown for different probe diameters. Easily it can be seen that both the curves for impedance and threshold value are similar. From Eq. (4) it follows that the skin behaves as a simple RCL oscillator with a natural frequency ν_p in Hz:

$$\nu_p = \frac{3}{2\pi a}\sqrt{\frac{\mu_1}{\rho}} \approx 72/a. \tag{5}$$

The Figs 1 and 2 show that the curves of threshold values and skin impedances are identical. Hence, it follows that the perception of the velocity amplitude of the skin is independent on the frequency within the frequency range considered. This is in accordance with the fact that the sensation of a stimulus is dependent on the first derivative of the skin deformation.

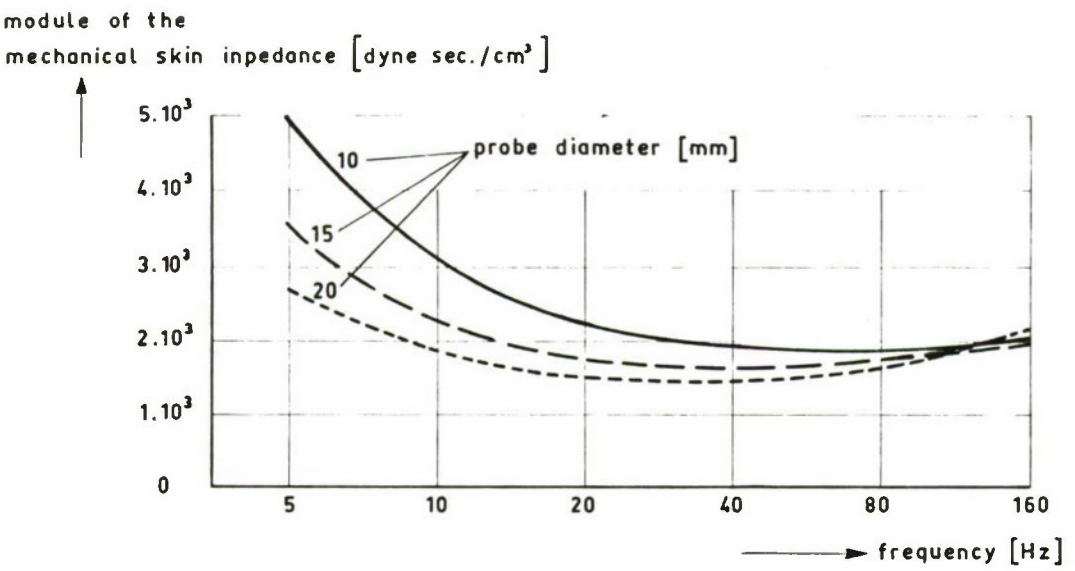


Fig. 2: Mechanical skin impedance as a function of frequency for three values of the probe diameter 2a.

Fig. 3 shows the results of experiments with the mechanical stimulator. The threshold values were determined for increasing as well as for decreasing amplitudes. The force applied for pressure on the skin was between 0.25N and 0.5N. The experiments were executed with three subjects. Four different positions of the probe with reference to the skin were applied. The threshold values are given in Fig. 3; they were 5 to 15 times higher than those found by Wilska [2]. This may be caused by the background noise in the room used, a phenomenon also described by Keidel [16]. Once more, the threshold values measured show that the perception of the stimuli is proportional to the first derivative of the displacement of the vibrator probe, hence to the skin velocity; a result which differs from that found in ref. 10.

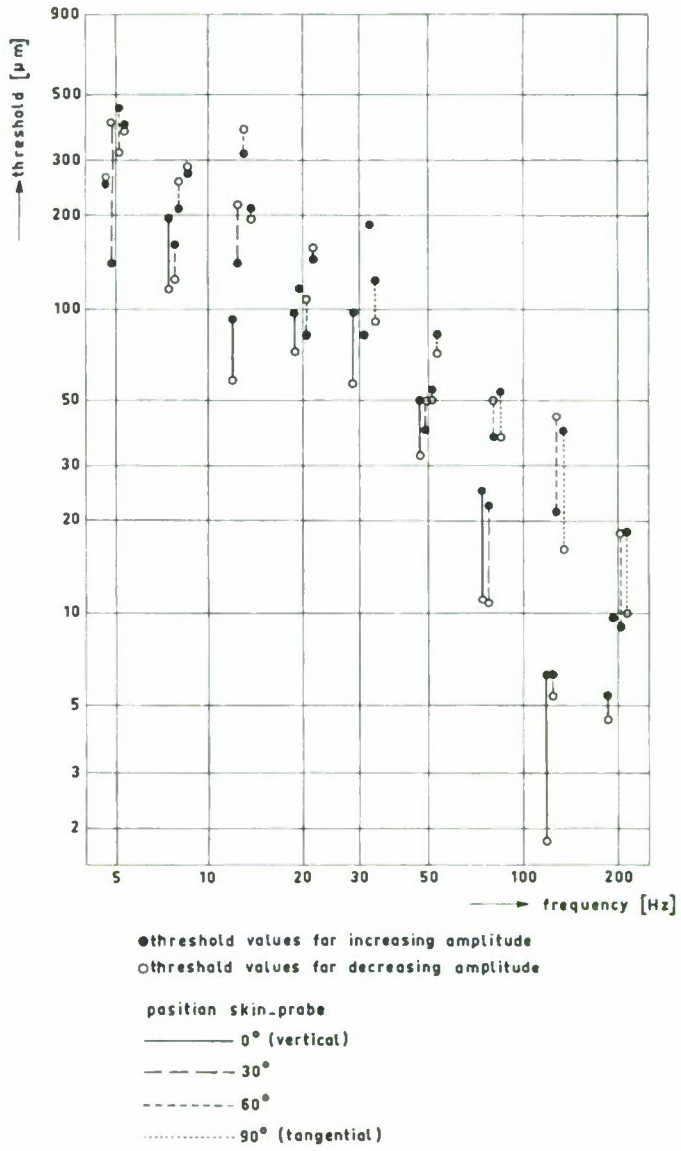


Figure 3: Threshold values of the just-perceptible amplitudes using the mechanical stimulator.

3. The just-noticeable difference

For pneumatic stimulation, the experiments concerned with the detection of the just-noticeable difference were based on the following items:

a. Only one stimulator had to be used. Preliminary experiments have shown that, if two stimulators were used, respectively by using both in time for different frequencies or by using sequentially one stimulator as reference and the other one to offer a certain frequency to the subject, the results were rather poor. Probably, the difference in sensation between left and right may also cause problems.
b. In using one stimulator, the time between two successive stimuli must be as short as possible.
c. The subject has not only to detect whether a difference exists between two successive stimuli, but must also indicate whether the frequency of the second stimulus is either equal to, or lower respectively higher than the first one.
d. To avoid variations in the psychological conditions in the course of the experiments, the whole testrun was programmed on magnetic tape, and thus standardized.

The difference between two frequencies was 0; 10; 15; 25 or 50%. The experiments have been executed for the center frequencies 14; 20; 37; 55; 80 and 103 Hz. A typical example for a center frequency of 80 Hz is: 80-40; 80-68; 80-100; 80-120; 80-72; 80-92; 80-88; 80-80 and 80-60. The amplitude used was far above the threshold value. The total program could be carried out within 20 minutes. The stimulator used had a diameter of 14 mm, and was mounted at the volar side of the left hand lower arm. In total 20 subjects were involved in this program.

The results are given in Fig. 4. This figure shows the percentage of correct decisions averaged over 20 subjects and calculated from the experiments for 20, 37, 55 and 80 Hz. Furthermore, the percentage of incorrect decisions is split up into two qualifications: class A being the case that "higher" instead of "lower" or "lower" instead of "higher" was indicated, and class B which is the situation that "equal" was mentioned instead of "higher" or "lower" or just the opposite. The curves show that discrimination within the interval of 15% is hardly possible. Moreover, if the rate of correct answers is observed for each center frequency separately, it turns out that discrimination for higher and lower center frequencies, such as 14 and 103 Hz, is difficult; a trend which is similar for the threshold values. In general, within a frequency range of 20 to 80 Hz the ability for discrimination is about 15%; outside this range, in particular at the higher frequencies, the discrimination is rather poor.

Similar experiments have been carried out with the mechanical stimulator. The ability for discrimination has been investigated for four center frequencies, viz. 5; 12.6; 32 and 80 Hz, and for five amplitudes, viz. 62.5;

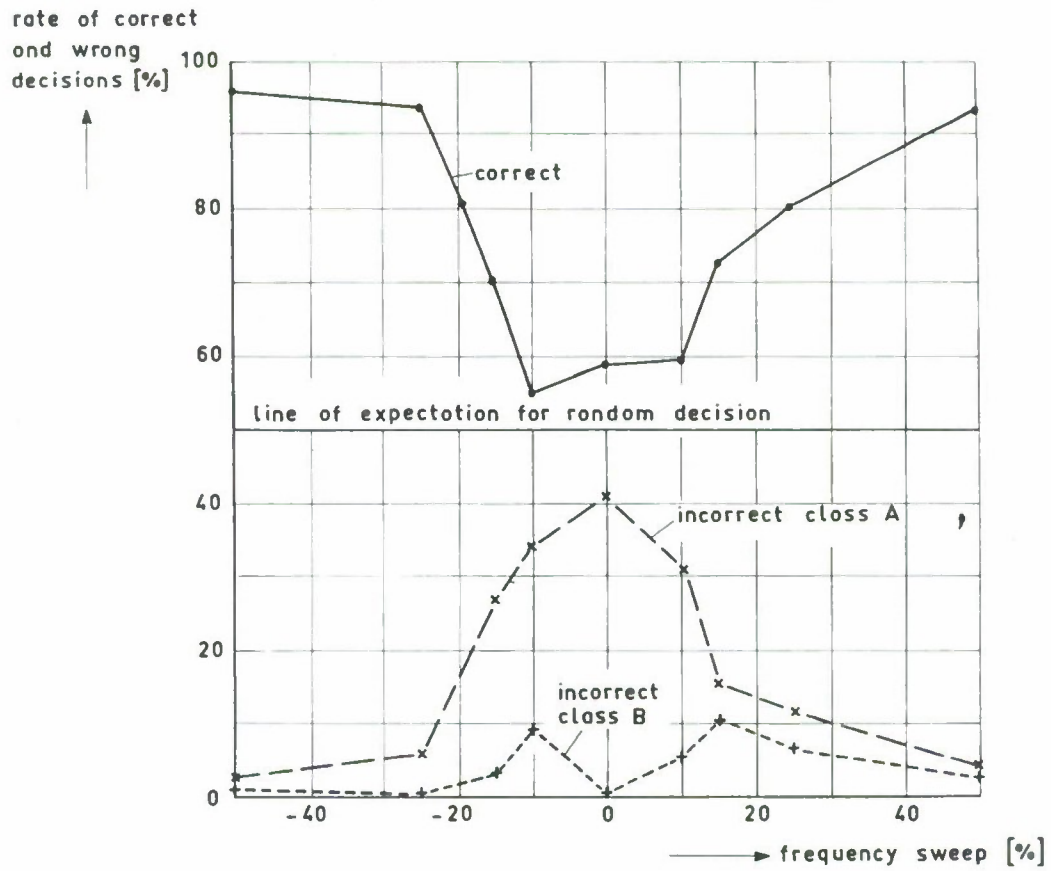


Fig. 4: Percentages "correct" and "incorrect" averaged over 20 subjects and over the center frequencies of 20, 37, 55 and 80 Hz.

125; 250; 500 and 1000 µm, while perpendicular as well as tangential stimulation has been applied. The task of the subject was here: After being offered a reference signal adjust another generator as close as possible to the reference generator. As an example the results of one subject are given in Fig. 5, and are indicating that a 15% discrimination for amplitude, and a 10% to 15% discrimination for frequency could be obtained.

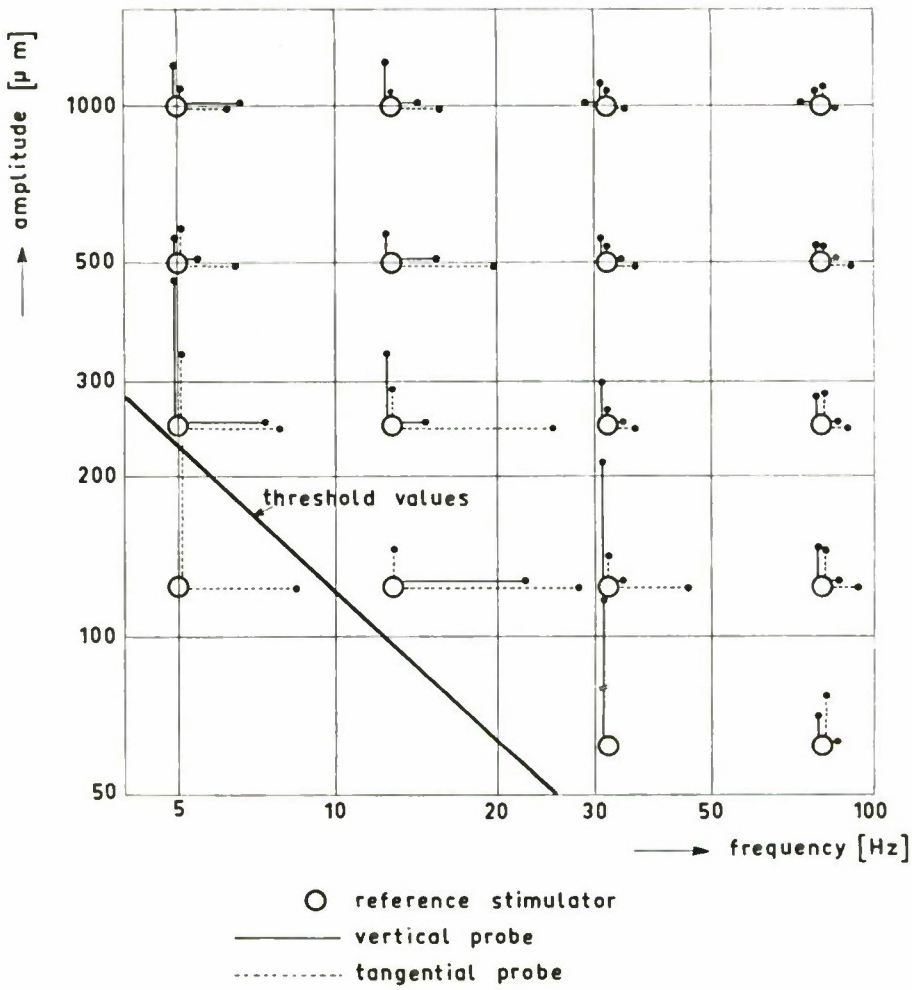


Figure 5: The just-noticeable difference in frequency and amplitude for a number of reference frequencies and amplitudes.

4. Tracking experiments

The applicability of tactile stimulation for the purpose of information transmission to the human operator was investigated in a simple manual tracking system. The experiments were focussed on answering the following questions:

1. In what way should the information be presented to the human operator, so that an optimal performance in the tracking task can be obtained? Hence, what will be better, an amplitude modulated signal or a frequency modulated one? Should the relation between the quantity to be presented and the output of the stimulator be linear or anti-logarithmic, in order to compensate for possible non-linearities such as Fechner's law? And, finally, in what way should the subject be given a zero reference?
2. How succesful is the optimal tactile information transfer system relative to comparable optimal visual or auditory display systems?

Fig. 6 gives a block diagram of the tracking system.

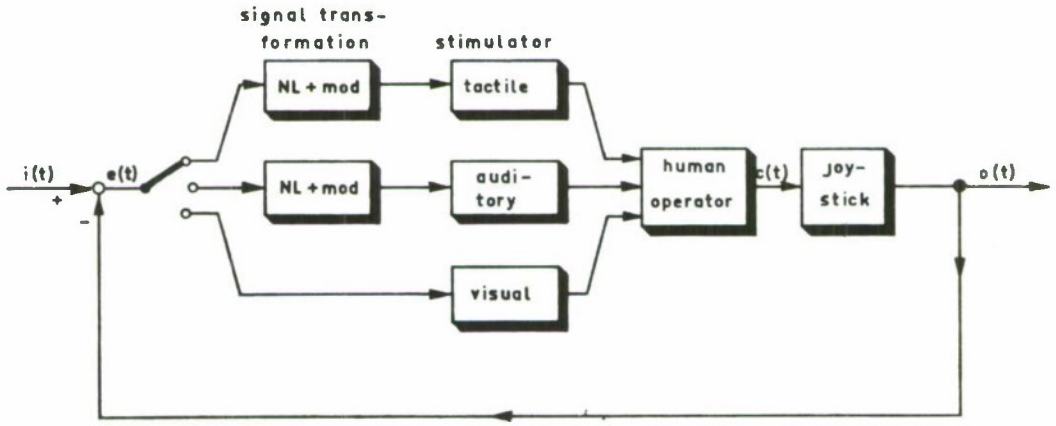


Figure 6: Block diagram of the tracking system.

The process to be controlled is a joy stick, a gain factor. The blocks indicated under the superscript "signal transformation" represent the relation between the error signal e(t) and the input of the stimulators applied. The types of signal transformation used in the experiments are presented in Table 1 for each of the modalities. The tactile experiments were performed using the mechanical stimulator, because the presentation of an amplitude modulated as well as a frequency modulated sinusoid was easier to accomplish with the mechanical stimulator than with the pneumatic device. In all tactile experiments the stimulator was fixed at the left hand side of the right hand shoulder as is shown in Fig. 7. As a consequence of the results of

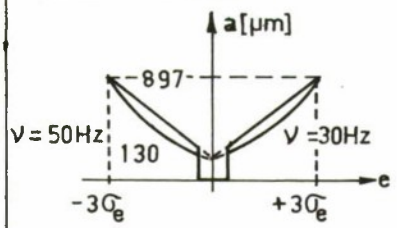
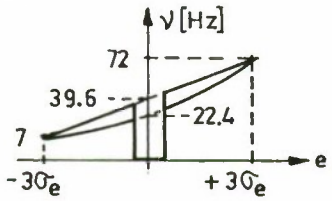
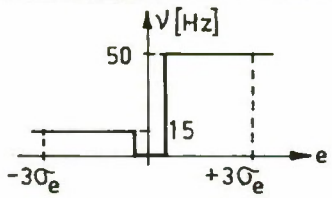
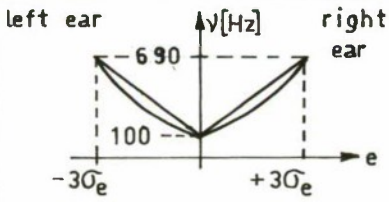
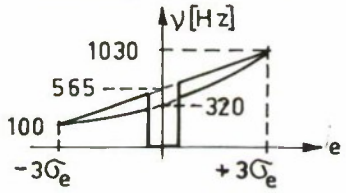
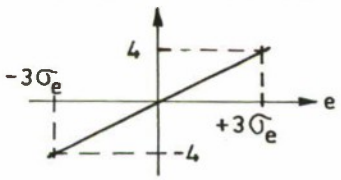
modality	display characteristics		type of modulation
tactile		linear or exponential, with and without dead zone; sign indication	AM
		amplitude: 1000 µm at 7 Hz 98 µm at 72 Hz	FM
		amplitude: 800 µm at 15 Hz 240 µm at 50 Hz 2 or 3 levels	two freq.
auditory		linear or exponential; no dead zone; sign indication	FM
		right ear only	FM
visual		moving horizontal line on CRT (displacement in cm)	none

Table 1: Types of presentation of the error signal e(t) investigated in the tracking experiments.

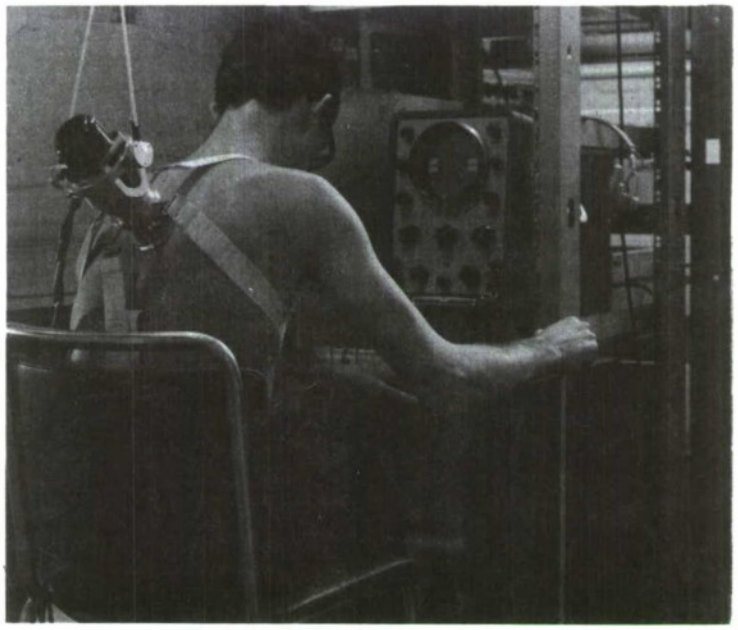


Figure 7: The set up for tactile tracking.

the threshold experiments in the frequency modulation experiments the amplitude of the carrier wave was varied inversely proportional to the frequency, hence the velocity amplitude remains constant.

For the auditory stimulation a pair of headphones were used having a flat amplitude characteristic in the frequency range considered. Frequency modulation was used, as being the most suitable way for auditory information transmission [17]. Here too, the velocity amplitude of the carrier wave was kept constant. The visual display was a four inch CRT with a horizontal line which moved in the vertical direction.

In all experiments the forcing function i(t) was white noise which was filtered by means of a 4th order band-pass filter between 0.02 and 0.1 Hz, and which was recorded on a magnetic tape. Two subjects, indicated in the following pages as S201 and S202, performed the experiments indicated in Table 1. The values of σ_e in the figures of this table necessary for the adjustment of the display characteristics were roughly estimated for each subject. The estimates were based on the outcomes of experiments measured at the end of a preliminary training period. Most of the experiments reported here had a duration of 5 minutes and were executed 2 times by each of the subjects. A few experiments have

been performed over a longer interval. In the latter case all relevant signals were recorded on magnetic tape and were used afterwards to determine the parameters in a mathematical model of the human operator. The results obtained concerning the model of the human operator are not treated here, but are given in Meyer's MSc-thesis (in Dutch) [18] and will be reported in the Annual Report 1970 of the Man-Machine Systems Group. During the experiments over a 5 minute interval there always was a period of 3 minutes in which the forcing function was exactly the same for all experiments. However, the starting point of the forcing function was always different to prevent the subject to notice this identity. In two experiments the magnetic tape was played back at double speed, hence the frequency range of the forcing function was 0.04 to 0.2 Hz. The subjects were also unaware of the fact that in some experiments the display had a linear characteristic and in others an exponential one. As a measure for the subjects performance the mean value of the error signal e(t),

$$a(t) = \frac{1}{T} \int_{t-T}^{t} e(\theta)d\theta, \tag{6}$$

and the RMS value of e(t)

$$b(t) = \frac{1}{T} \int_{t-T}^{t} e^2(\theta)d\theta, \tag{7}$$

were determined ($T = 180$ sec, $t \geq 180$ sec).
To normalize the quantities a(t) and b(t) they were divided by σ_i, which is the RMS value of the forcing function i(t) measured over the same 3 minute observation period. The results of the experiments are summarized in Table 2. In this table the quantities a' and b' are the values of a/σ_i and b/σ_i respectively averaged over 3 experiments; the quantity $\Delta b'$ is the maximum deviation from the mean value of b' calculated from the 3 values of b/σ_i. The quantity B' is the estimate of b' determined at the end of the training period, and thus a measure for the quantity σ_e in Table 1 (σ_e was used to adjust the display characteristics). The dead zones are expressed in percentages of the range: $-3\sigma_i$ to $+3\sigma_i$.

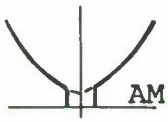
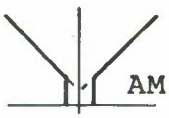
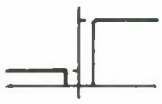
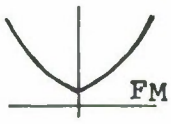
display characteristics	nr	dead zone [%]	subject 201 B' [%]	b' [%]	Δb' [%]	a' [%]	subject 202 B' [%]	b' [%]	Δb' [%]	a' [%]	
 AM	1	0	70	59	7	0	35	26	7	0	tactile
	2	2,5	60	54	3	30					
	3	5	40	35	3	15	30	23	4	0	
	4*	5	70	63	2	21					
	5	10	35	29	4	15	35	26	2	15	
 AM	6	5	40	34	3	0	30	19	2	0	
	7*	5					55	45	4	15	
FM	8	0	90	88	1	30	85	78	2	30	
	9	2,5	60	50	3	0	65	62	2	0	
	10	5	50	42	2	0	50	41	2	0	
	11	10	50	41	1	0	40	35	1	0	
 2 or 3-state	12	0		33	3	15		28	3	15	
	13	2,5		15	1	0		18	1	0	
	14	5		17	1	0		19	2	0	
	15	10		26	1	0		24	1	0	
 FM	16	0	20	13	1	0	20	13	1	0	audotory
FM	17	0	20	15	4	0	20	11	1	0	
	18	2,5	20	11	1	0	20	14	1	0	
	19	5	20	12	1	0	20	15	1	0	
	20	10	30	22	1	0	30	25	1	0	
	21	0	15	11	2	0	15	8	3	0	visual

Table 2: Results of the tracking experiments for tactile, auditory and visual information presentation.
*: forcing function in the range: 0.04 - 0.2 Hz.
In all other cases: 0.02 - 0.1 Hz.

5. Results

The results are summarized in Table 2, in which the experiments are labeled by the number 1 up to and including 21. The Figs. 8 and 9 show some typical records of the tracking behavior of the subjects. For the tactile stimulation the following information can be deduced from the Table 2.

a. Frequency modulation (nrs 8; 9; 10; 11) is a type of presentation which yields worse results than amplitude modulation (nrs 1; 2; 3; 4; 5). This might be explained by the fact that for frequency modulation the stimuli (see Table 1) are rather close to the threshold values (see Fig. 3), while in this region also the just-noticeable differences are rather high (see Fig. 5).

b. The amplitude modulation presentation is also not very successful. Although a zero reference is given by means of a sign indication, both the subjects introduce a test signal (see Figs 8 and 9). Probably, this test signal is used in order to get information on the deviation of the error from the zero reference. This suggests that only the information on the sign of the error signal is taken into account.

c. The results of the two-state presentaion (nr 12) are not worse than those of the experiments with amplitude modulation. This affirms the assumption that the subjects only use the sign of the error signal. The Figs 8 and 9 also illustrate that the nature of the output signal of the subjects in the case of amplitude modulation is identical to that of the two-state presentation.

d. On the one hand, the introduction of a dead zone about the zero reference leads to a loss of information, on the other hand, the subjects are better informed about the order of magnitude of the deviations. Hence, it may be expected that an optimal dead zone can be found. The results show very clearly that the introduction of a dead zone provides a much better performance (compare nrs 1; 8; 12 with nrs 3; 10; 14). The optimal width of the dead zone is about 5% of the range of the signal.

e. By comparing the exponential display with the linear one (nrs 3 and 6) it follows that the most important difference is the suppression of the test signal for the linear display (see Fig. 9). That for an exponential display the subject still needs to generate a test signal may be explained by the fact that signal amplitudes just beyond the dead zone are rather close to the threshold value.

f. Stimulation with a display having a three-state presentation results in the lowest values of the errors. In this case, the subjects do not introduce test signals

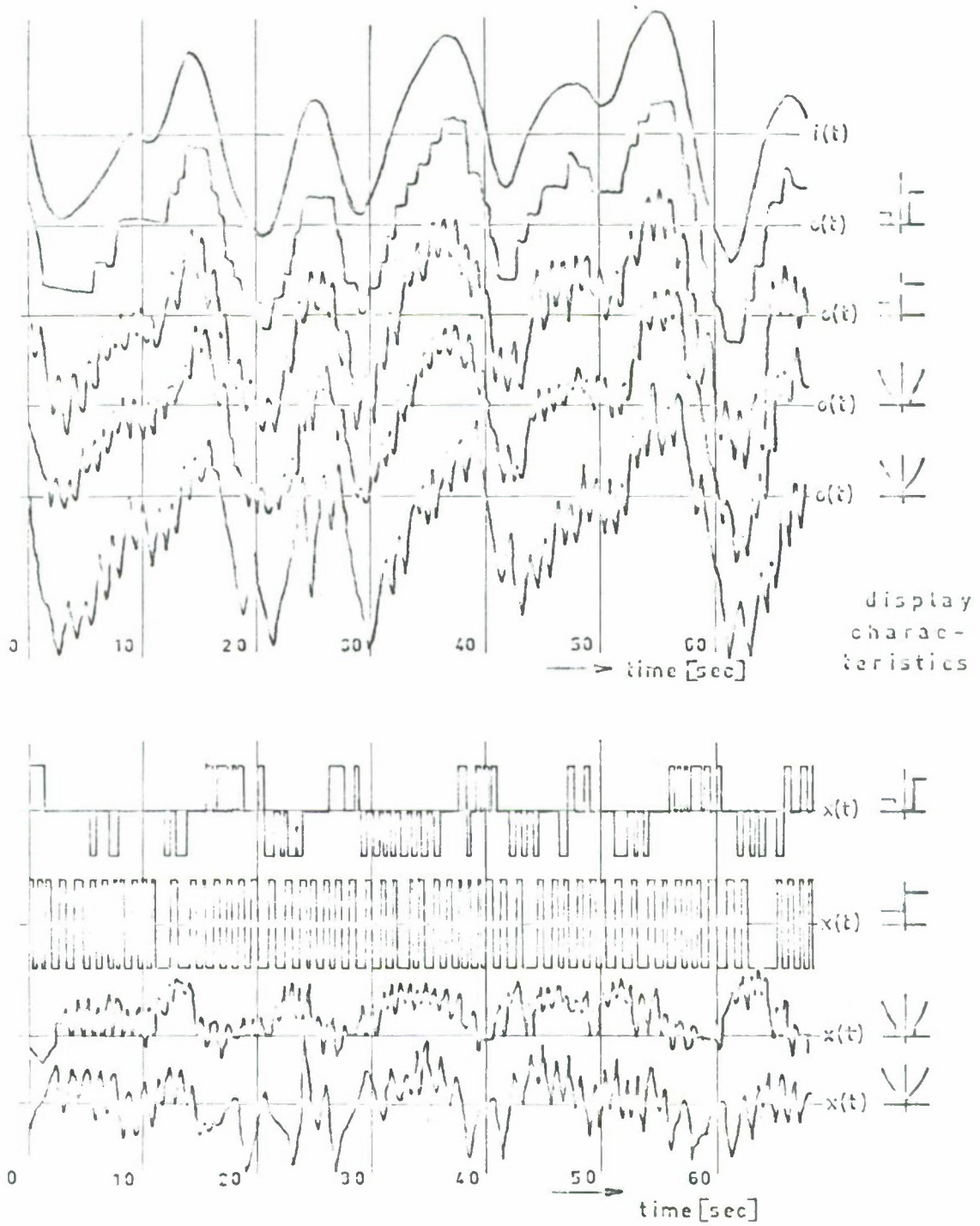


Figure 8: The forcing function i(t), the outputs of the joy stick o(t) and the error signals x(t) after the non-linear element but before the modulator, for some of the tactile experiments executed by subject 201. Dead zone if applied: 5%.

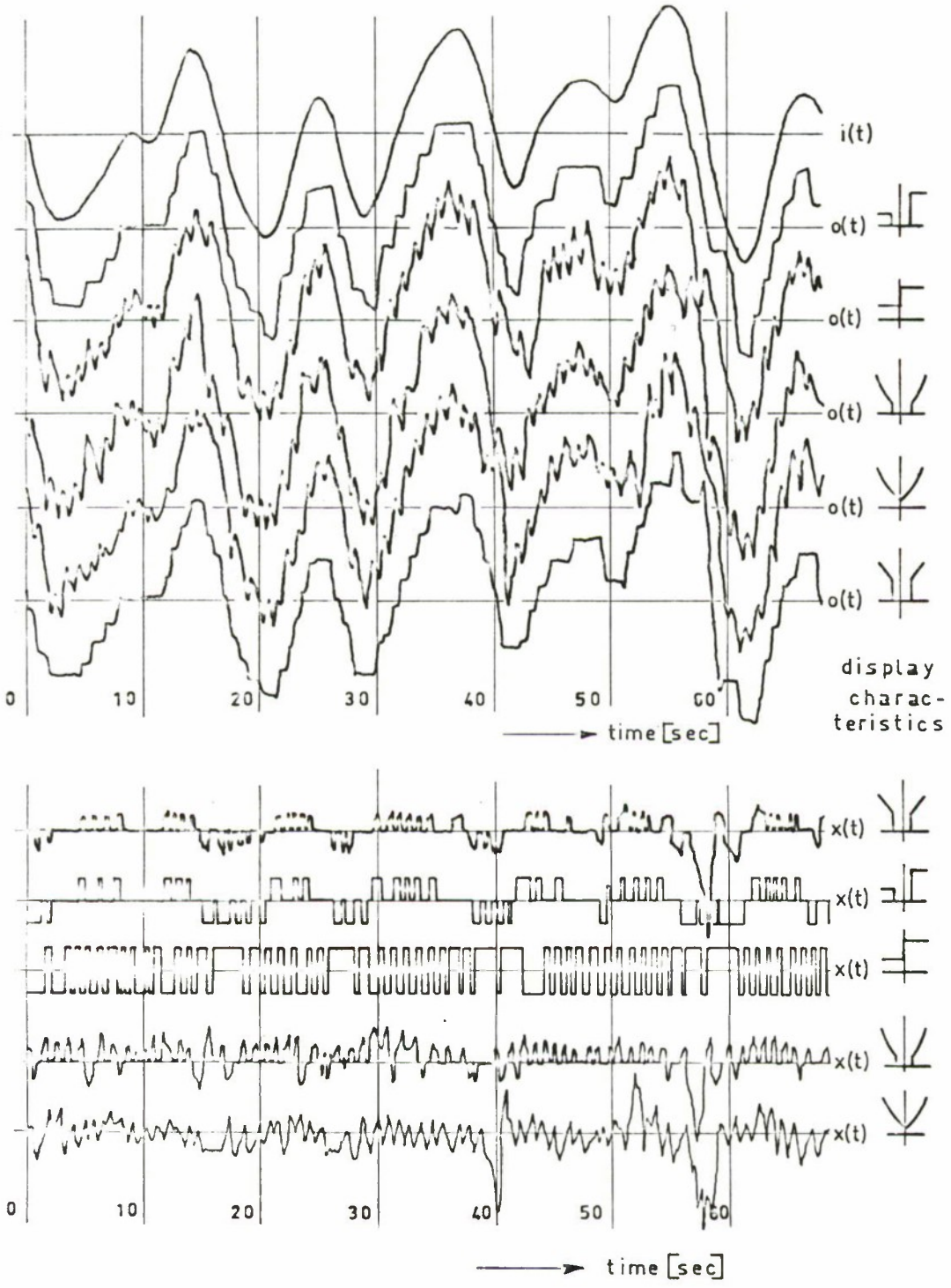


Figure 9: The forcing function i(t), the outputs of the joy stick θ(t) and the error signals x(t) after the non-linear element but before the modulator, for some of the tactile experiments executed by subject 202. Dead zone if applied: 5%

(see Figs 8 and 9); they only act when a stimulus is percieved.

g. If the frequency range of the forcing function i(t) is changed from $0.02 < \nu < 0.1$ Hz to $0.04 < \nu < 0.2$ Hz, the relative error increases by a factor 2. Here it should be noted that the subjects probably were limited in the control of the joy stick due to human operator output dynamics.

Finally, a comparison between tactile, visual and auditory tracking will be made.
The following items are of interest.

a. The application of a dead zone of a width of 2.5 and 5% also for auditory stimulation leads to a better performance.

b. The results obtained from the experiments with the optimal visual and auditory presentation on the one hand, and the tactile three-state presentation with a dead zone of 5% on the other hand, show that the mutual differences are rather small.

c. A very interesting fact is that the subjects remarked that the execution of the control task using tactile stimulation was less tiring than in the case of visual or auditory stimulation.

To get an idea about the influence of boredom, fatique and inattention of the subjects during the tactile tracking task, some of the most successful experiments of Table 1 have been executed for an observation time of 12 and 60 minutes. Due to the fact that the frequency modulated presentation was inferior, only the amplitude modulation has been considered. The results represented by the relative error b/σ_i for a total observation time of 12 minutes are given in Fig. 10. The figure shows that in in particular the three-state presentation leads to a fairly stationary behavior of the subject. The rather high values of b/σ_i for S202 in the time interval of 9 to 11 minutes were caused by inattention of the subject during the eighth minute (remember that b/σ_i is averaged over 3 minutes). The two-state presentation always results in the introduction of test signals (see the Figs 8 and 9), hence, in general the factor b/σ_i will increase. The introduction of a dead zone reduces the relative error very strongly. The Fig. 11 shows the quantity b/σ_i as a function of the observation time for a total time of 60 minutes. The figure shows that during the first 15 minutes the behavior of the subject is fairly stationary; hereafter he became due to the uncomfortable sitting position during the experiments. The enormous increase in b/σ_i at the moment $t = 17$; 25; 35 and 41 minutes are due to the fact that the subject changes his sitting position during which he was almost inactive. At the moment $t = 50$ minutes the subject was told that the end of the test was in sight; this fact probably influenced his behavior in a positive sense.

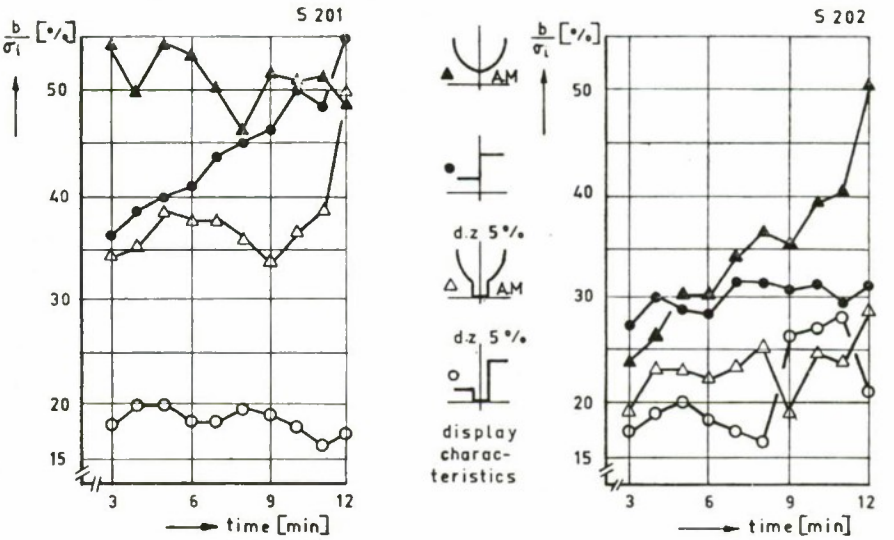


Figure 10: The relative error b/σ_i of both the subjects as function of time for different types of stimulation; total observation time 12 minutes.

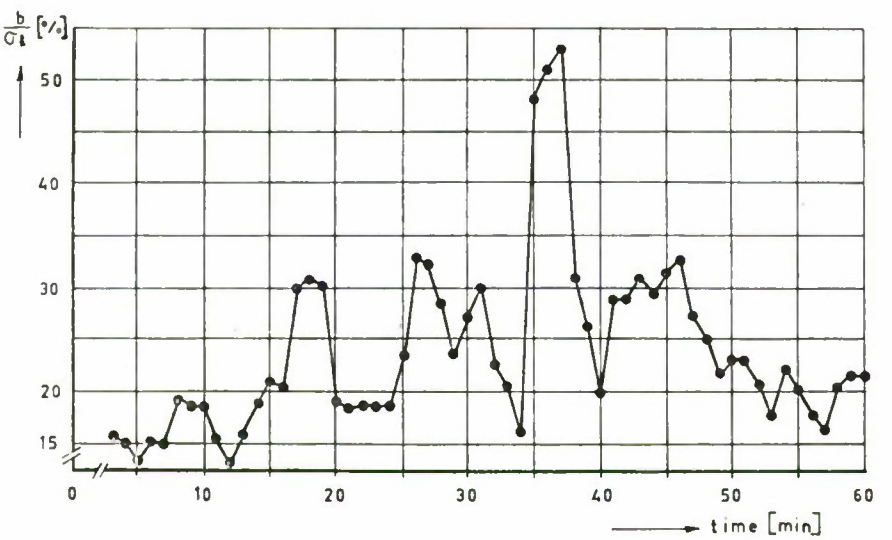


Figure 11: The relative error b/σ_i of subject S202 as a function of time for the three-state tactile display; total observation time 60 minutes.

6. Conclusions

The experiments reported suggest that information processing via the skin for use in prosthesis control will be possible with only one stimulator for one variable, for instance the grasping force.

The stimulator output should be in an amplitude range of 0.1 to 1.0 mm, and in a frequency range of 5 to 80 Hz.

The three-state presention gives the best results in a continuous tracking task, a result which is very attractive because this type of stimulation can easily be realized in prosthesis control. Moreover, it is experienced by the subjects that the mental load is lower for the tactile stimulation than for the visual and auditory stimulation.

7. Acknowledgements

The authors feel greatly indebted to Mr W.H.S. Mandl for his valuable contribution in this investigation, in particular, with respect to the development of and the experiments with the mechanical stimulator. Likewise, the authors wish to express their gratitude to Mr W. Verkaik for his contribution in the investigation with the pneumatic stimulator.

8. References

1. Lunteren, A. van; Stassen, H.G.: Annual Report 1969 of the Man-Machine Systems Group, Laboratory for Measurement and Control, Dept of Mech. Eng., Delft University of Technology, Delft (1970), Ch V.
2. Wilska, A.: On the vibrational sensitivity in different regions of the body surface, Acta Physiol. Scand. (31) (1954), pp. 285-289.
3. Hugony, A.: Uber die Empfindung von Schwingungen mittels des Tastsinnes, Z. Biol. 96 (1935), p. 584.
4. Verillo, R.T.: Investigation of some parameters of the cutaneous threshold for vibration, J. Acoust. Soc. Am. Vol. 34 (1962), pp. 1768-1173.
5. Békésy, G. von: Can we feel the nervous discharges of the end organs during vibratory stimulation of the skin. J. Acoust. Soc. Am. Vol. 34 (1962), pp. 850-856.

6. Keidel. W.D.: Vibrations reception, Verlag: Universitätsbund, Erlangen (1956).
7. Eijkman, E.G.J.; Vendrick, A.J.H.: Dynamics of the vibration sense at low frequencies, J. Acoust. Soc. Am. Vol. 32 (1960), p. 1134.
8. Goff, R.: In: Cutaneous channels for communications, Sensory communications, Rosenblith, W.A., MIT Press (1961).
9. Anderson; Attell, B.; Munson, W.A.: Electrical excitation of nerves in the skin at audiofrequencies, J. Acoust. Soc. Am. Vol. 23 (1951), pp. 155-159.
10. Verillo, R.T.: Effect of contactor area on the vibrotactile threshold, J. Acoust. Soc. Am. Vol. 35 (1963), pp. 1962-1966.
11. Fechner, G.T.: Elemente der Psychophysik, Leipzig (1862), pp. 548-560.
12. Stevens, S.S.: The psychophysics of sensory function, Sensory Communication, Rosenblith, W.A., MIT Press (1961).
13. Bliss, J.C.; Linvill, J.G.: A direct translation reading aid, reading alphabetic shapes tactually, St. Dunstan's Int. Conf. on sensory devices for the blind.
14. Oestreicher, H.L.: Field and impedance of an oscillating sphere in a viscoelastic medium with an application to biophysics, J. Acoust. Soc. Am. Vol. 23 (1951), pp. 707-714.
15. Gierke, H.E. von; Oestreicher, H.L.; Franke, E.K.; Parrack, H.O.; Wittern, W.W. von: Physics of vibrations in living tissues, J. Appl. Physiol.(4) (1952).
16. Keidel, W.D.; Spring, M.: Neurophysiological evidence for Stevens powerfunction in man, J. Acoust. Soc. Am. Vol. 38 (1965), p. 191.
17. Shower, E.G.; Biddulph, R.: Differential pitch sensitivity of the ear, J. Acoust. Soc. Am. Vol. 3 (1931) no 2.
18. Meyer, A.W.A.: Het gedrag van de mens in een volgsysteem met tactiele informatie overdracht. (The behavior of a human operator in a tactile tracking task), Laboratory for Measurement and Control, Dept. of Mech. Eng., Delft University of Technology, report A-74 (1970).

Can proprioceptive cues unload the human operator ?

Karl-Friedrich Kraiss[+)]

Abstract

It's a well known fact, that the manual control of a dynamic system can strongly be influenced by the control force, which is perceived at the control stick. In spite of several hints in the literature, there are only few directions for the design of optimal control-forces.

Manual control movements require the action of an internal proprioceptive control-loop and an external visually supervised control-loop as well. Perceptions resulting from the control-force at the stick are, by definition, only suited to modify the behaviour of the proprioceptive control loop.

To determine the influence of various types of control force (spring, viscous damping, inertia), experiments were made concerning a step tracking movement of the arm in a horizontal plane. Included were 20 male subjects, 19-21 years old.

It is shown by the experiments, that the shape of the control movements is significantly modified by the type and amount of control-force. Optimal selected control-forces result in a lengthening of the proprioceptive controlled interval, simultaneously shortening the visually controlled part of the movement by approximately 20 [%]. The visual channel of the human operator is discharged remarkably, herby justifying the expense which is necessary to simulate suitable control-forces.

+) Forschungsinstitut für Anthropotechnik, 5309 Meckenheim, Germany
This paper is part of a more comprehensive report entitled "Zum Einfluß des Bewegungswiderstandes auf die Güte manueller Regelung". Institutsbericht No.4

1. Introduction

The output of information from a human operator into a control-loop mainly consists of the movement of his arms. As to the pilot, this action is the manual deflection of the control stick, hereby overcoming the control resistance. While moving the manipulator, the pilot perceives a control feel, which he associates to the behaviour of his airplane according to his exercise and experience. The intended output to the machine therefore results in an additional input, which is reflected from the machine to the pilot (fig. 1). This feedback is a very essential part of the input information. In special flight situations with restricted flexibility of the visual channel (low level flight) the control feel seems to be indispensable.

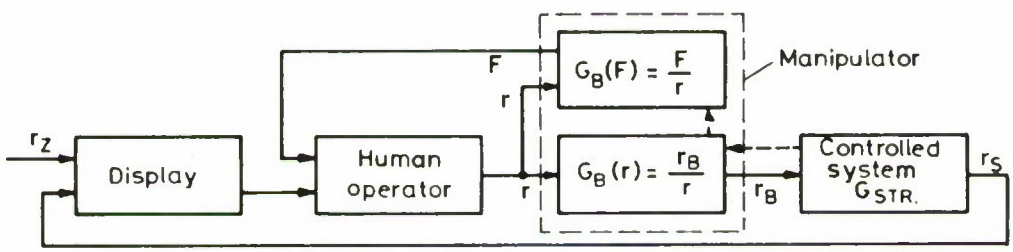


Figure 1. Control force F as additional information-feedback to the human operator.

Though the importance of appropriate control-forces has been recognized very early [1], there are only few recommendations for the layout of a suitable control feel. Numerous investigations with special control-force configurations didn't lead to a better understanding of the perception and processing of proprioceptive information by the human operator [2 ./. 9]. Therefore it seems to be necessary to continue the research in this field, starting with a more detailed analysis of intended movements.

2. The coordination of movements

The definition of general directions for control-force tailoring presumes, that we know in which way the shape of a movement is influenced by various

control forces and control systems. To answer this question, we start with a description of the mechanisms, which are the basis for a control movement.

If we define a control action as an intentionally performed movement, this is only a coarse description of the process going on. In each movement, even in the simplest one, a lot of muscle pairs are involved. The movement is initiated by the force developed of one muscle against his antagonist. It's abvious, that these mechanical details are not conciously controlled by the human operator. The pilot does only know the time he wants to start his movement and the target he wants to reach. In addition to that, he has an idea of the velocity he should achieve to fulfill his task [10].

Therefore a rather complex control programme is necessary, to coordinate the activity of various muscle-pairs in the desired way. Figure 2 presents a hypothetic scheme of a programme-controller for this purpose, as derived from physiological research [10].

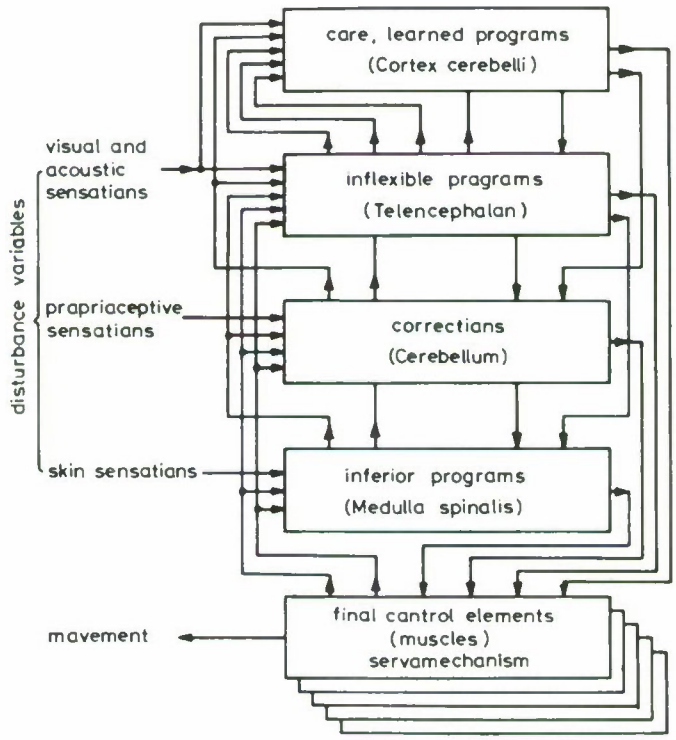


Figure 2. Programme-controller for manual movements (modified from Ranke [10])

As we learn from this, there are different levels of consciousness in the hierarchy of the brain and manifold interconnections. The servomechanisms at the end of the chain translate the nervous signals to the intended movements. It should be emphasized, that according to this scheme, proprioceptive cues are feed back to the cerrebellum, while visual informations reach the brain on a higher level (Telencephalon). That means, that there are two different control loops. Control movements result from the cooperation of the internal proprioceptive and the external visually supervised control-loop as well. Tschaidse [11] presumes, that the internal loop controls the automated details, while the external loop governs the conscious part of a movement. From the different time constants of the visual [120 ms] and proprioceptive [40 ms] loop results a nonlinear system. Figure 3 illustrates the two control-loops of the human operator involved in a pursuit tracking task.

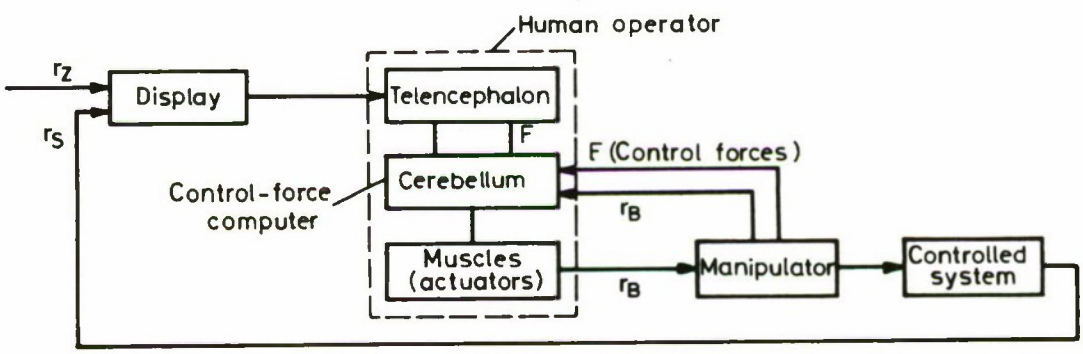
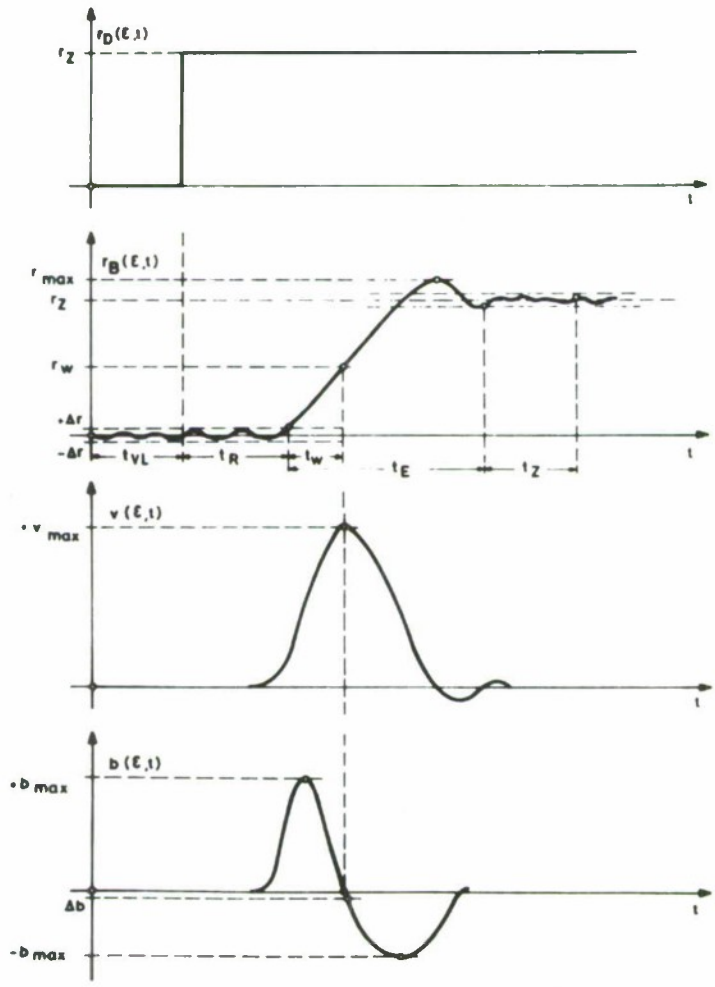


Figure 3. The visual and proprioceptive control-loop of the human operator.

According to these considerations, a well learned positioning movement consists of an internally governed quick part at the beginning, that is followed by a visually controlled interval when reaching the target. The approach to the target is connected with a slow-down of the velocity. To analyse the duration of both intervals a step tracking movement can be used. In this special pattern, the switching from proprioceptive to visual control-mechanism is indicated by the turning point of the curve, i.e. the point of maximum velocity. Beginning with this time, the velocity decreases continuously. A schematic diagramm of a manual step tracking function is presented in figure 4. Velocity and acceleration curves are included.

The position of the turning point (indicated by t_w, r_w in fig. 4) can be influenced by several variables as f.e. the desired target accuracy [12]. But there is nothing known about the effect of control-forces F perceived at the manipulator (fig 3). According to the considerations above we can assume, that the length of the proprioceptive controlled interval indicates the effectiveness of additional informations taken from the perceived control forces. If the total adjusting time t_E (see fig. 4) remains constant, lengthening of t_w means shortening of the visual supervised interval. By this effect the visual channel of the human operator is unloaded. Starting form this point of view, the best control-force configuration should yield the longest values for the proprioceptively controlled interval.



ε	step direction
$R(\varepsilon,t)$:	unit step desired value
$r(\varepsilon,t)$:	step response
$v(\varepsilon,t)$:	step velocity
$b(\varepsilon,t)$:	step acceleration
Δr :	target accuracy
t_R :	reaction time
t_E :	adjusting time
t_Z :	time in target
t_W :	time till turning point
t_{VL} :	lead time

Figure 4. Schematic diagramme of a step tracking function.

The course and amount of stick deflection, that is necessary to pursuit a step is determined by the characteristics of the controlled system. In contrast to that, the perceived control force depends on the mechanical properties of the stick. To demonstrate these relations, figure 5 presents idealized pursuit movements for controlled systems of 0., 1. and 2. order. The corresponding courses of the control force F are dash-lined for the three basic types of resistance :

$F = K \cdot r$: spring tension

$F = D \cdot r^{\bullet}$: viscous damping

$F = M \cdot r^{\bullet\bullet}$: inertia

As we learn from figure 5, the course of the control force can become rather complicated ever for the simple task of step tracking.

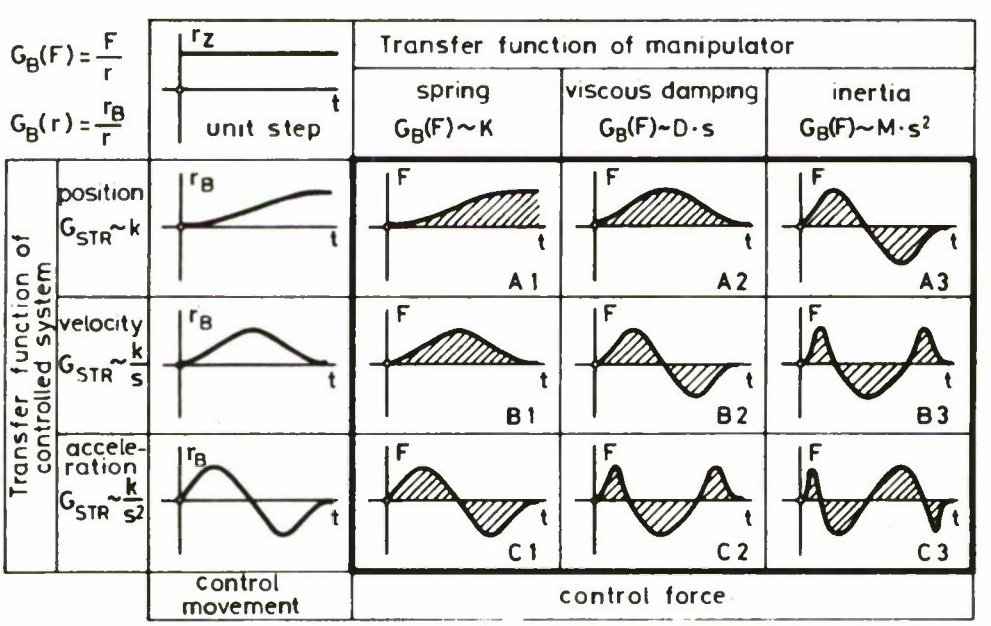


Figure 5. Schematic control movements and corresponding control-forces for a step tracking task with respect to various types of controlled systems and control-force.

3. Experiments

The purpose of the described experiments is to determine the shape of a well learned step tracking curve and the variations arising with various types of control-force. The experiments are performed for movements of the arm in a

horizontal plane. The experimental configuration is presented schematically in figure 6.

task : To pursuit a target, that is jumping pseudo-randomly in 8 different directions (ϵ_D) starting from the center. Instruction: "as quick as possible".

The pursuit variable should remain at least 0.3 sec within the targetregion (see t_z, fig. 4). The radius of the target-circle (Δr) is 0.6 [cm]. The step amplitude is 12 [cm].

system : To restrict the experimental programme, only a positioning system is used (see diagrammes A in fig. 5).

manipulator : stick, 70 cm long.

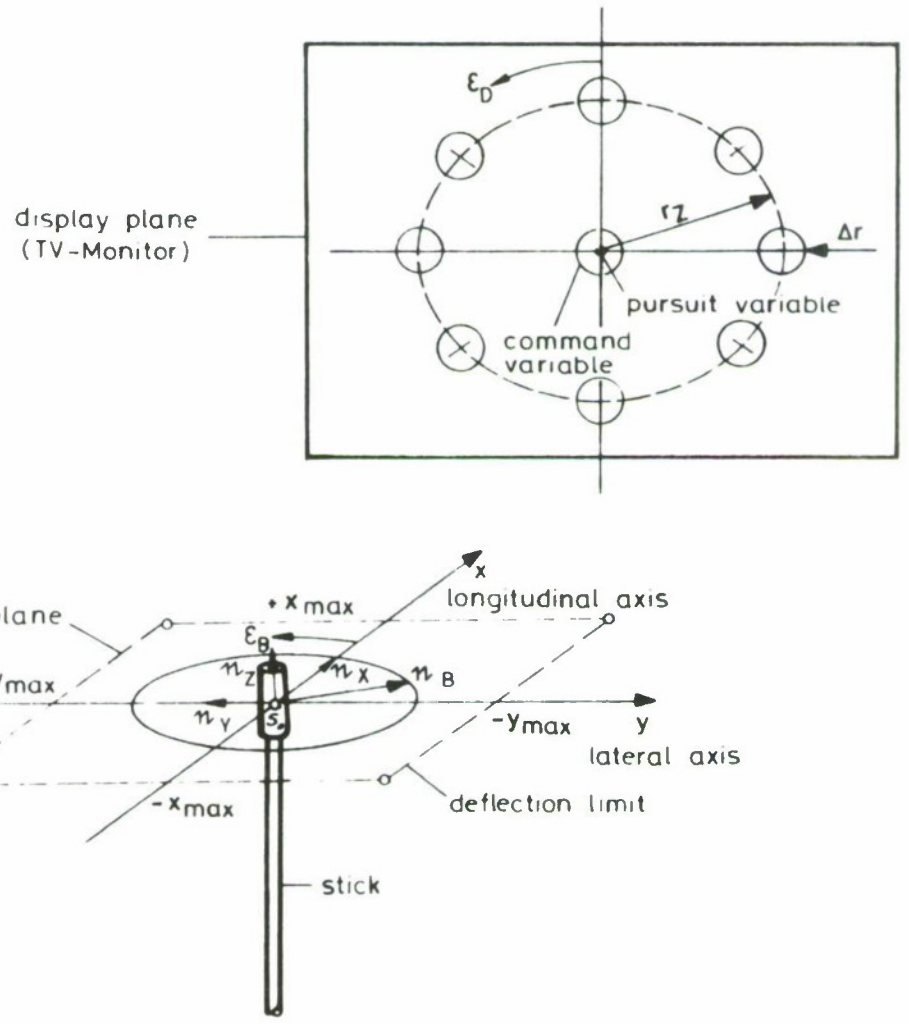


Figure 6. Experimental configuration

control force : Various parameters of spring-tension, viscous damping and inertia can be simulated at the stick by electro-hydraulic means.

experimental programme : Figure 7 shows the 56 different testconditions, which are presented to each subject seven times in a pseudo-random manner.

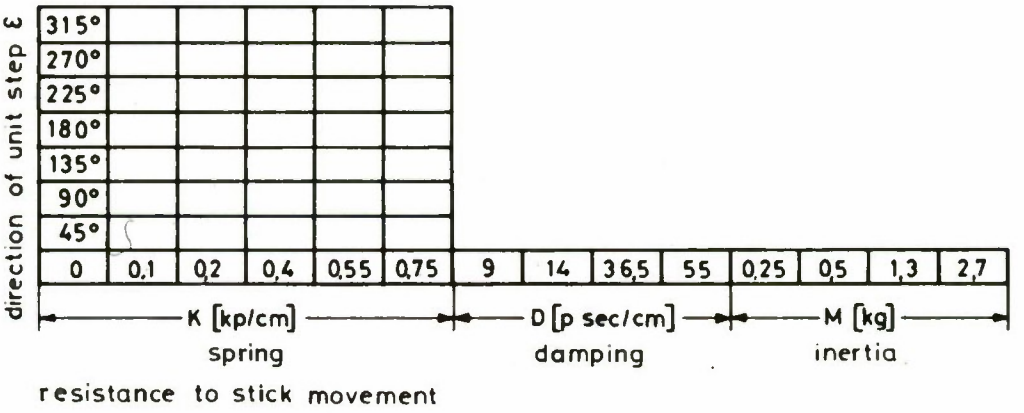


Figure 7. Matrix of test-conditions.

subjects : 20 male persons, 19 - 21 years old, mixed education, right-handed.

Data processing : The experimental results are based on the analysis of 8.000 manual step tracking actions. Each test condition is measured 140 times. The data have a gaussian-distribution at a significance-level of 99 [%]. The standard deviation is 8,5 [%] related to the mean value.

Environment : Test chamber 2 x 3 x 2 m, sound attenuated, low level illumination, temperature ca. 20° C.

4. Results

The experimental results are summarized in figure 8. It is shown, that the parameters of the step tracking curve are modified significantly, when spring-tension (spring-constant K) is applied to the stick. In contrast to that, there are almost no variations in the case of viscous damping (coefficient D) and inertia (coefficient M).

If we analyse the movement-structure, the position of the turning point (t_w, r_w) is especially important. Therefore figure 9 presents the diagramms $t_w = t_w$ (K) and $r_w = r_w$ (K) with enlarged scale factor. It is shown, that the position r_w of the turning-point between start and target is described by the ratio:

$$r_v : r_p = 1 : 2 \quad \text{for} \quad K = 0 \text{ [kp/cm]}$$

If the spring-constant increases, the turning point moves toward the target, reaching his maximum for K = 0,15 [kp/cm]. At this point the ratio has changed to :

$$r_v : r_p = 1 : 4$$

For further increasing spring constant K, the turning point removes continuously from the target, hereby enlarging the visual controlled interval of the movement.

The diagramme $t_w = t_w$ (K) in figure 9 shows, that the adjusting time t_E consists of a proprioceptive (t_p) and a visually (t_v) controlled time-interval. Obviously t_v can be shortened by introducing a spring-constant of about 0.1 [kp/cm] to the stick. Dependent on the amount of spring-tension, the following parts of the total adjusting time t_E have to be supervised visually (see fig. 9) :

K = 0 [kp/cm] : t_v = 64 [%]
K = 0,1 [kp/cm] : t_v = 50 [%]
K = 0,4 [kp/cm] : t_v = 67 [%]

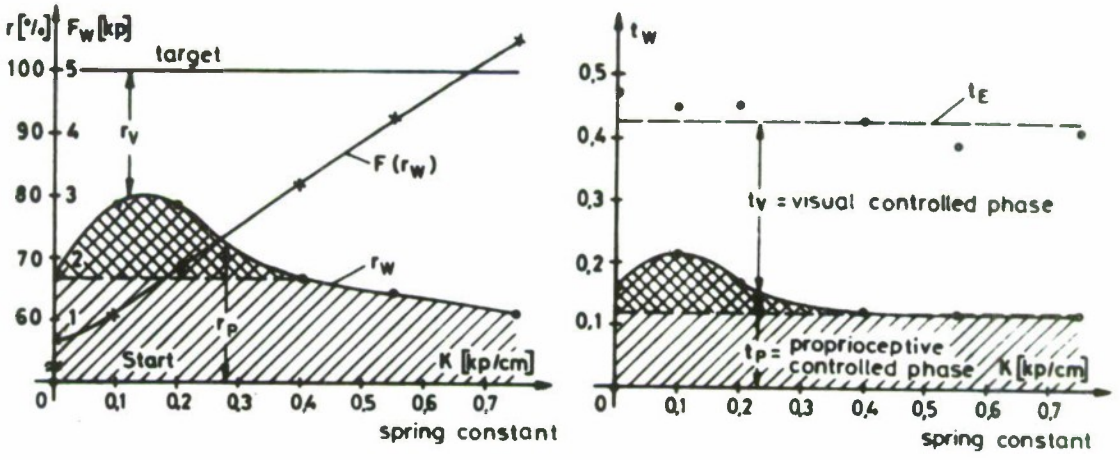


Figure 9. Shortening of the visually controlled interval t_v of a manual step action by proprioceptive cues. (For the symbols see fig. 4, fig. 8).

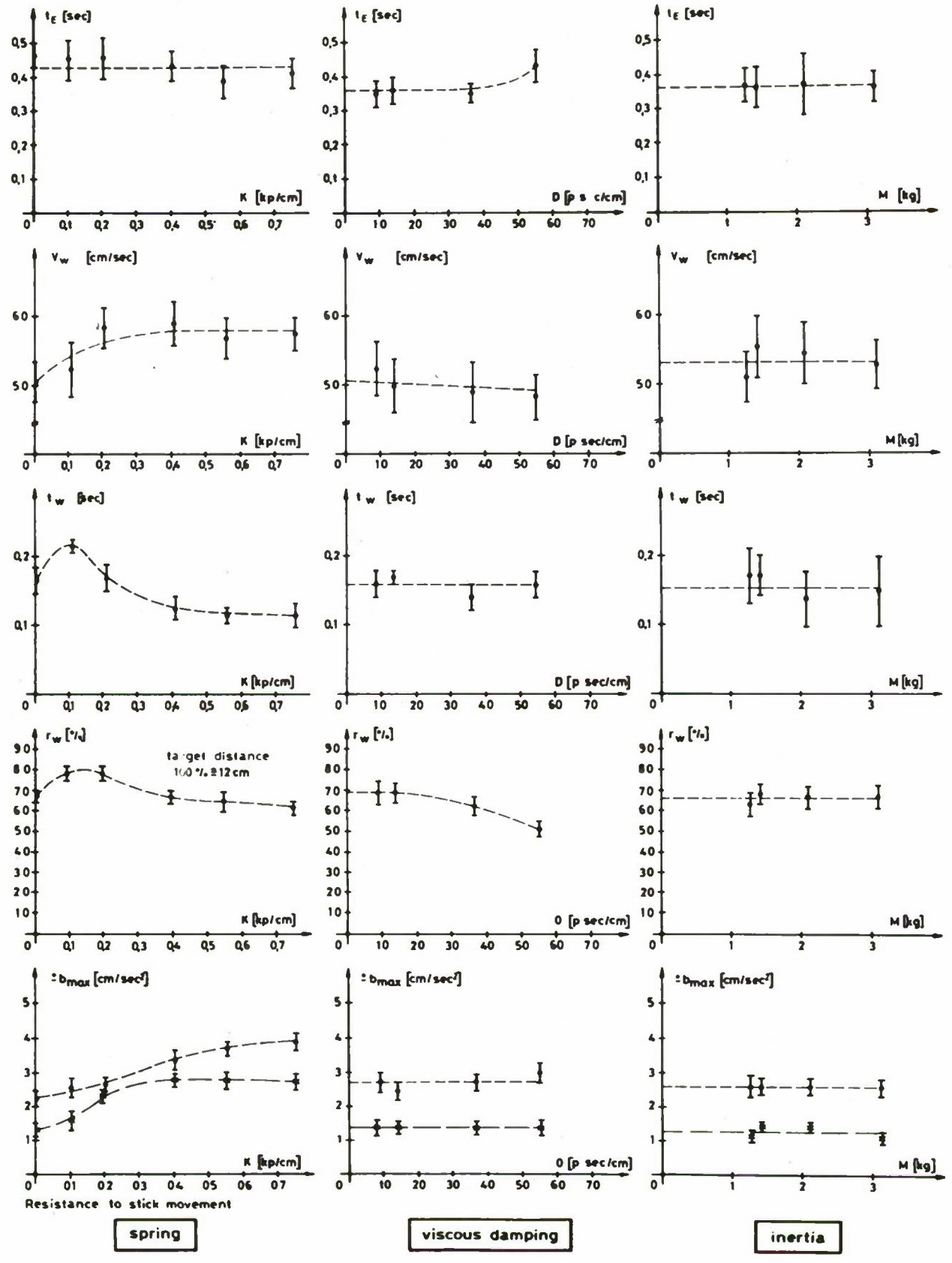


Figure 8. The shape of a step tracking action for different types of control force (for the symbols see fig. 4)

These results confirm the hypotheses an the influence of additional proprioceptive informations discussed in paragraph 2 of this paper :

> For the control of a positioning system, a spring-restrained stick should be used. The spring-constant should be 0.15 [kp/cm]. The effect of proprioceptive cues from the perceived control-force consists in a shortening of the visual supervised control-interval by approximately 20 [%]. Hereby the visual channel of the human operator is remarkably discharged and free for other tasks.

The described experiments included only a positioning system. To continue this research, similar experiments should be performed with typical systems of higher order as f.e. velocity- and acceleration-systems. In addition to that, complex systems have to be examined, to determine in which way the different proprioceptive channels cooperate.

These rather extensive experiments may yield the results that are necessary to formulate general directions for the layout of suitable control forces.

(1) Orlansky, J.

The human Factor in the Design of Stick and Rudder Controls

Aeronaut. Engng. Rev., 1949, 8, 1-10

(2) Abramovitz, M.
Schmidt, S.F.
Van Dyke, R.D.

Investigation of the Use of a Stick Force proportional to pitching Acceleration for normal Acceleration Warning

NACA RM A 53E21 Aug. 53

(3) Mallick, D.L.
Reeder, J.P.

Flight Evaluation of several Spring Force Gradients and Bobweights in the Cyclic-Power-Control System of a light Helicopter

NASA TN D 537

(4) Faber, S.
Crane, H.L.

A longitudinal Control Feel System for in-flight Research on response Feel

NASA TN D 632 Jan. 61

(5) Watler, J.F.

Effects of static Friction and Force Displacement Gradient on a discrete visual Tracking Task

Long Beach, Calif., Douglas Aircraft Company Report LB 31050, 1963

(6) Graham, D.

Research on the Effect of Nonlinearities on Tracking Performance

AMRL-TR-67-9, July 1967
Wright Patterson Air Force Base, Ohio

(7) Wasicko, R.J.
Magdaleno, R.E.

Effects of Nonlinearities on Human Operator Tracking Performance: A Review of the Literatur

AMRL-TR-65-158, Okt. 1965
Wright Patterson Air Force Base, Ohio

(8) McRuer, D.T.
Magdaleno, R.E.

Human Pilot Dynamics with various Manupulators

Technical Report AFFDL-TR-66-138, Dec. 1966
Wright Patterson Air Force Base, Ohio

(9) Magdaleno, R.E.
Mc. Ruer, D.T.

Effects of Manipulator Restraints on Human Operator Performance

Technical Report AFFDL-TR-66-72, Dec. 66
Wright Patterson Air Force Base, Ohio

(10) Baader, E.W.

Handbuch der gesamten Arbeitsmedizin
Band 1 : Arbeitsphysiologie

Urban & Schwarzenberg
Berlin, München , Wien, 1961

(11) Tschaidse, L.W.

Die Koordinierung willkürlicher Bewegungen beim Menschen im Blickwinkel allgemeiner Gesetzmäßigkeiten der Steuerung und der Steuersysteme

In Probleme der Kybernetik,
Vol. 8, 1965 S. 334-371

(12) Aldrich, J.W.
et. al.

A formal Model for Arm Motion during Target Approach

5th Annual Nasa-University Conference on Manual Control MIT,
March 27-29, 1969

DESIGN OF CORRECTIVE DYNAMICS CONTROLLERS

Ralph Mekel
Department of Electrical Engineering
The City College of The City University of New York

Patrick Peruo, Jr.
Navigation and Control Division
Bendix Aerospace-Electronics Co. Teterboro, N. J.

ABSTRACT

This informal paper represents a continuing effort to enable the engineer to apply Liapunov's stability theory to the design of high order nonlinear control systems. The paper develops a design technique for a class of nonlinear controllers which stabilize high order unstable nonlinear control systems and bound the system's settling time within prescribed specifications. The design technique is based upon a class of Liapunov functions and applies to systems with multiple nonlinearities. The advantage of this design technique over several other existing techniques lies in the fact that one can generate the form of the Liapunov function by inspecting the system matrix. Furthermore, the quantitative properties of the class of Liapunov functions are fully utilized by using a design criterion function which is formulated from the negative ratio of a Liapunov function and its time derivative.

The purpose of this design technique is to study nonlinear man-machine control system modeling. One of the main objectives of the study is to realize a controller which will provide corrective dynamics to improve an assumed (or tentatively formulated) human operator mathematical model in a closed loop model-reference control system to compensate for nonlinear and adaptive behavior of an actual human operator. The realized controller becomes part of the human operator's mathematical model.

After establishing that the mathematical model of the human operator is representative of a "normal" actual human operator, one can study the usefulness of the mathematical model for investigating changes in the actual human operator due to changes in environment and due to different tasks. These changes can then be interpreted in terms of the mathematical model's nonlinear coefficient variations.

In order to introduce the design technique let the nonlinear augmented control system be represented by Fig. 1

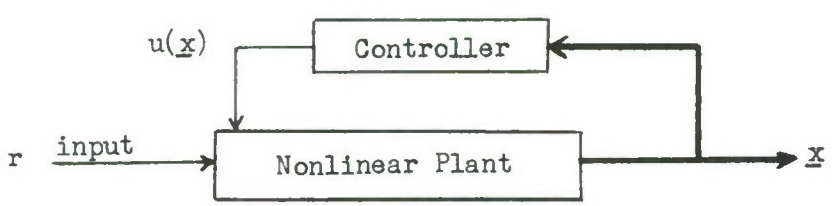


Fig. 1. Nonlinear Augmented Control System Block Diagram

and described by the nonlinear differential equation

$$\dot{\underline{x}} = A_o\underline{x} + A_{on}(\underline{x})\underline{x} \tag{1}$$

In Fig. 1 the nonlinear plant is known and the controller is to be designed. In Eq. (1) A_o is a constant square matrix that characterizes the linear part of the augmented system. Matrix $A_{on}(\underline{x})$ is also a square matrix which characterizes the nonlinear part of the augmented system. The subscripts o and n denote the order of the system and the number of nonlinearities in the system respectively; $\underline{x}$ denotes a column matrix of the state variables and $\dot{\underline{x}} = d\underline{x}/dt$.

The next step in the design procedure is the formulation of the class of Liapunov functions, their time derivatives and the design criterion function in the forms

$$V(\underline{x}) = \underline{x}'M_o\underline{x} + \underline{x}'M_{on}(\underline{x})\underline{x} > 0 \tag{2}$$

$$\dot{V}(\underline{x}) = \underline{x}'D_o\underline{x} + \underline{x}'D_{on}(\underline{x})\underline{x} < 0 \tag{3}$$

and

$$\eta(\underline{x}) = -\frac{\dot{V}(\underline{x})}{V(\underline{x})} \tag{4}$$

where M_o and D_o are square matrices with constant elements. Matrices $M_{on}(\underline{x})$ and $D_{on}(\underline{x})$ are also square matrices whose elements consist of the nonlinear functions of the state variables and their derivatives and are directly related to those of the $A_{on}(\underline{x})$ matrix. The crux of this development is the formulation of the $M_{on}(\underline{x})$ matrix which can be obtained in a routine way from the system matrix $A_{on}(\underline{x})$. Note that the prime denotes the transpose.

At this state of the design one has to satisfy Liapunov's stability criterion for asymptotic stability. One way of satisfying this requirement is to constrain the elements of the D_o and $D_{on}(\underline{x})$ matrices, denoted by d_{ii} and d_{ij}, to satisfy the following conditions

$$d_{ii} < 0 \tag{5}$$

and

$$d_{ij} + d_{ji} = 0 \tag{6}$$

Equations (5) and (6) produce a set of inequalities which permit the evaluation of the unknown elements in the D_o and $D_{on}(\underline{x})$ matrices. Note that by determining the elements of the D_o and $D_{on}(\underline{x})$ matrices one also determines the elements of the M_o and $M_{on}(\underline{x})$ matrices at the same time. Sylvester's theorem may be applied to check the positive and negative definitness of these matrices.

The set of inequalities obtained from Eqs. (5) & (6) and the design criterion function given by Eq. (4) contain all the information necessary for the design of the corrective dynamics controller. A digital computer is now utilized to solve the set of inequalities. When employing the digital computer as an aid in the design one does not have to rely heavily neither

on intuition nor on a large random parameter search. It is the systematic and logical use of the computer program that assures one with an effective and economic design of the corrective dynamics controller.

This design technique has been utilized to realize a controller which provides corrective dynamics to improve an assumed (or tentatively formulated) human operator mathematical model in a closed loop model-reference control system to compensate for nonlinear and adaptive behavior of an actual human operator. The model-reference system and the realized controller are shown in Fig. 2.

For more detailed information the reader is referred to the following publications:

1. Mekel, R.: A Class of Liapunov Functions for High Order Nonlinear Control Systems. Proceedings of the Asilomar Conference on Circuits and Systems, Monterey, California, Nov. 1-3, 1967.

2. Mekel, R.: Synthesis of Corrective Dynamics for Human Operator Models in Closed-Loop Control Systems. NASA, Langley Research Center, LWP-816, October 21, 1969.

3. Peruo, P.,Jr.: Design of Nonlinear Control Systems Via A Time Domain Ratio Criterion. Ph.D. Dissertation, The City University of New York, New York, N. Y. 1970.

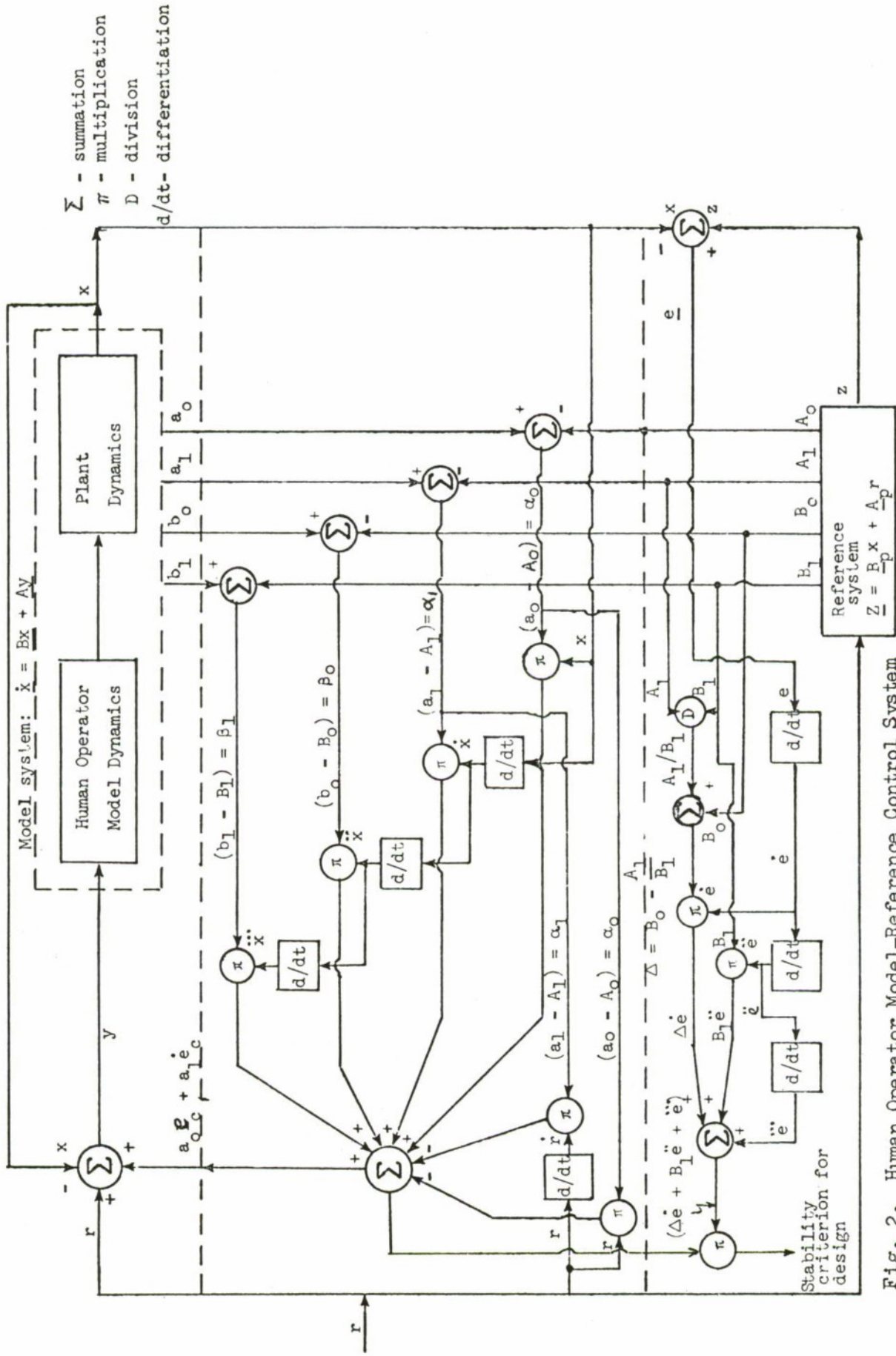


Fig. 2. Human Operator Model-Reference Control System with Corrective Dynamics Controller.

A TWO DIMENSIONAL, PATTERN RECOGNIZING, ADAPTIVE MODEL OF A HUMAN CONTROLLER

D. W. Gilstad and K. S. Fu
Purdue University

Abstract

A two dimensional adaptive model of a human operator in a visual-manual compensatory tracking task is proposed. The model is adaptive to changes in gain and/or form of the plant dynamics on each axis, accommodating plant transfer functions of the form (k/s) or (k/s^2). Pattern recognition is used to sense changes in the control system requiring a change in the model operating mode. The form of the model is derived from consideration of the physiological processes evident in a human controller. It contains separate sections dealing with perception of the control system variables, action taken by the central nervous system as a result of the sensory inputs, and the conversion of CNS commands into motions of the controller. The model reflects the influence of the dominant, type 2 control strategy on the type 1 strategy when tracking is performed with heterogeneous plant axis dynamics. Tests of the gain adaptive algorithm simulating time invariant human operator performance are described, and the test results presented for type 1 and type 2 control situations. Results of pattern recognition tests conducted with model-generated data are also presented.

Introduction and Background

The problem of modeling a human operator in various control situations has received a number of different treatments (1),(2),(3). Depending on their intended degree of generality, the linear and quasi-linear models proposed range from relatively simple to complex describing functions. These functions have largely been derived with the aid of frequency domain or statistical analysis. Some models include random noise injection to represent that part of the operator's output which is not linearly related to the input signal. Time domain analysis has also been used, though less widely, to develop models. It is the method which was used in this study.

It involves the choice, by some method not specified, of a suitable function to represent the human operator. The function may be linear or nonlinear and is provided only in its basic form. The adjustment of the parameters of this form in such a way as to optimize a given rule (function) for each one is the task to be performed. One may build a system with parameter adjustment capability inherent in it and run this system in parallel with a human operator to facilitate adjustment. With the addition of performance criteria, we have a model matching system (4). Figure 1 is a block diagram of a single axis model matching system; the second

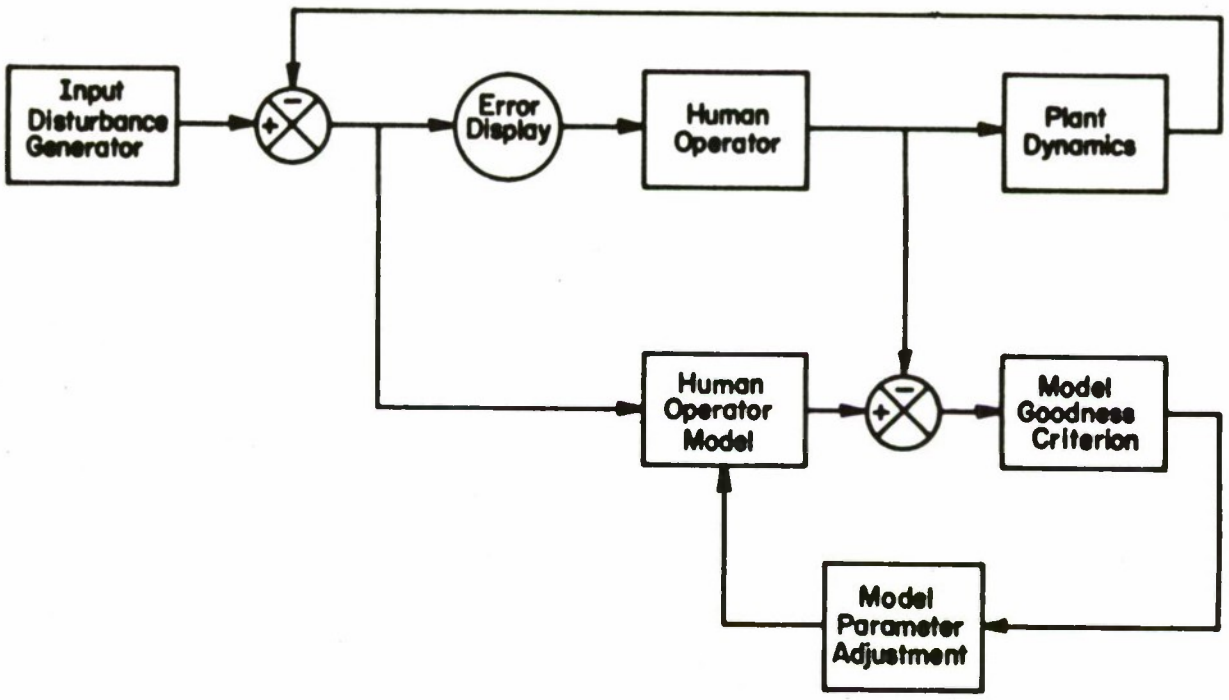


Figure 1. Model Matching System.

control loop is deleted for clarity. Note that the model goodness criterion, although shown applied to the controller output, could as well be applied to the system output by including plant dynamics after the model in the lower loop and changing the signal pickoff point in the upper loop. Model matching was used in this study; the model goodness criteria were applied by a human observer, and model parameter adjustments were accomplished through the use of an interactive mode on a high speed digital computer.

The form of the model proposed herein is an extension of the single axis model of Fu and Gould (5) to two dimensions. A portion of the gain calculation algorithm furnishes gain adaptability, and mode switching, based on pattern recognition, allows the accommodation of two types of plant dynamics. The choice of the model form was governed by a desire to simulate, as closely as possible, the physiological processes evident in a human controller. Figure 2 shows the structure of each of the two single axis sections. The four blocks between the error display and the plant dynamics were modeled separately to allow inclusion of the available pieces of physiological information. When more understanding is gained regarding any or all of the human controller subsystems, the appropriate section(s) of the model may be modified within the framework given.

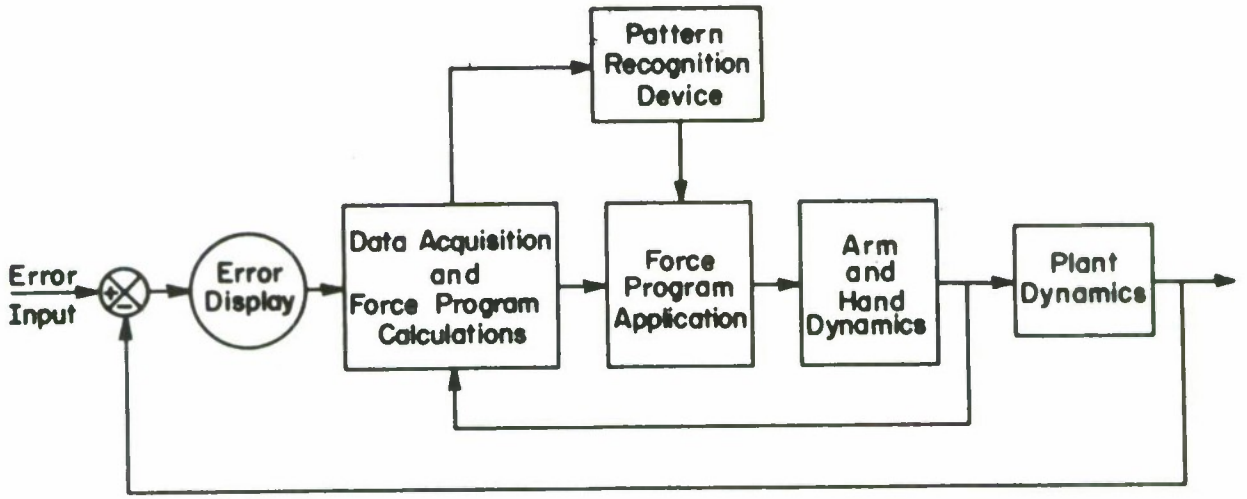


Figure 2. Single Axis Adaptive System.

Experimental Equipment and Procedures

An experimental program was designed to provide data which could subsequently be analyzed to aid in synthesizing the desired model. The experiments were conceived to gather information on the performance of a human operator performing a compensatory, two axis, visual-manual tracking task with the control system dynamics changing at discrete intervals of time. Figure 3 is a block diagram of the control system considered. The method of error presentation was entirely visual; control actions and error inputs were simultaneously displayed on the face of an oscilloscope and positioned with reference to an illuminated reticle which covered the display area. A single lighted dot approximately 1 mm in diameter appeared at the center of the reticle when no error or control inputs were present. During each run of the experiment, the operator attempted to keep the dot in the center of the reticle by applying control inputs which would null the error inputs. Keeping the dot precisely in the center was, of course, impossible since the operator was given no indication of the error inputs or control system outputs other than the position of the single dot. He had to allow the dot to move off-center in order to decide what movement of the controller was required to recenter it. The magnitude of the operator's error was the radial distance of the dot from the center at any time; the angular position of the dot was not considered in determining the quality of the operator's performance.

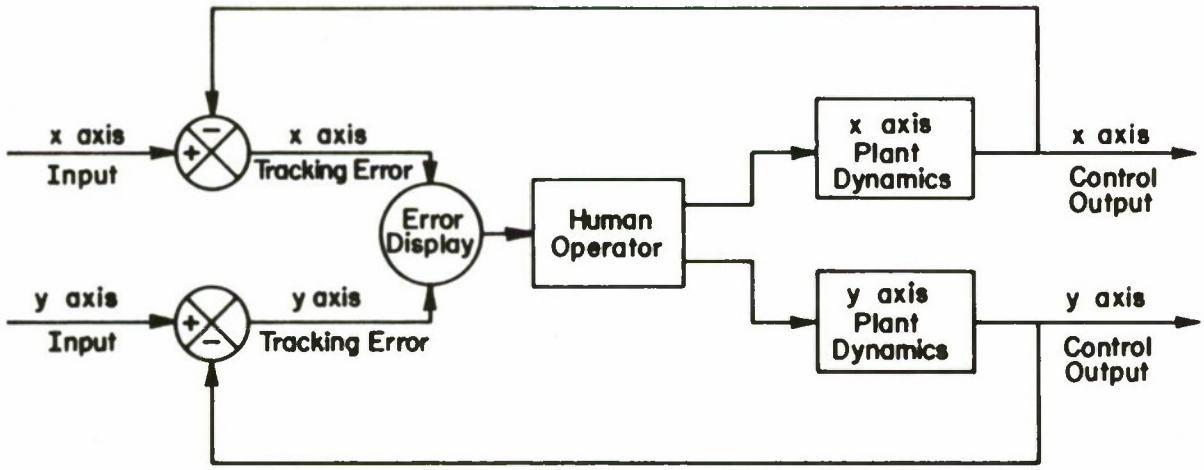


Figure 3. Two Loop Manual Control System.

The error signal used in the experiments was a sum of sinusoids in the frequency range 0.01 to 0.10 hertz. The object was to provide random appearing smoothly varying signals which simulated the error inputs to a manual control system in some real environment. It was inportant to prevent the operator from anticipating the strength and phase of the error signal and tracking it precognitively; by making use of any discernable pattern in the error signal, the operator would have more information at his disposal than the display was meant to provide. Although precognitive tracking behavior is important in some manual control systems, the experiment was designed to eliminate this specific kind of behavior. Therefore, some effort was directed toward insuring that the magnitude and phase of the error sensed by the human controller were totally unpredictable.

Plant dynamics were simulated on the analog portion of a hybrid computer. The two forms considered had transfer functions as follows

$$G(s) = k/s \quad \text{(single integration - type 1)}$$

$$G(s) = k/s^2 \quad \text{(double integration - type 2)}$$

Pertinent data from the tracking tests, including the input error components, controller and plant outputs, and their resultants, were converted to digital data and stored by the digital section

of the hybrid computer. The tracking runs, whose timing was controlled by the digital clock, consisted of the following sequence of events.

50 seconds of tracking with homogeneous type 1 plants, followed immediately by

50 seconds of tracking with homogeneous type 2 plants, followed by a

15 second rest period for the operator, then

50 seconds of tracking with homogeneous type 1 plants, followed immediately by

50 seconds of tracking with a type 2 plant on the x-axis and a type 1 plant on the y-axis followed immediately by

50 seconds of tracking with homogeneous type 2 plants, followed by a

15 second rest period for the operator, then

50 seconds of tracking with homogeneous type 1 plants, followed immediately by

50 seconds of tracking with a type 1 plant on the x-axis and a type 2 plant on the y-axis, followed immediately by

50 seconds of tracking with homogeneous type 2 plants.

Periods of 50 seconds were chosen because the transition regions appeared to reach a steady state within them, and they were long enough to prevent the operator from accurately predicting the exact time of transition from one plant configuration to another. These transitions were smoothed by causing the integrating capacitor involved to track its input as an initial condition prior to putting its associated amplifier into operation. Dynamics transitions did not result in a discontinuous jump of the dot on the oscilloscope face, although, depending on the controller position at that time, smooth, rapid increases in the tracking error were sometimes observed.

The Adapted Model

Figure 4 is a block diagram of the adapted or single mode model. A model of this form is used for each of the two plant dynamics types considered in this study. The force program calculation block contains an algorithm which results in gain adaptive behavior. Data perception and preparation as viewed in the model are shown in greater detail in Figure 5. Deadband and saturation characteristic parameters were determined individually for

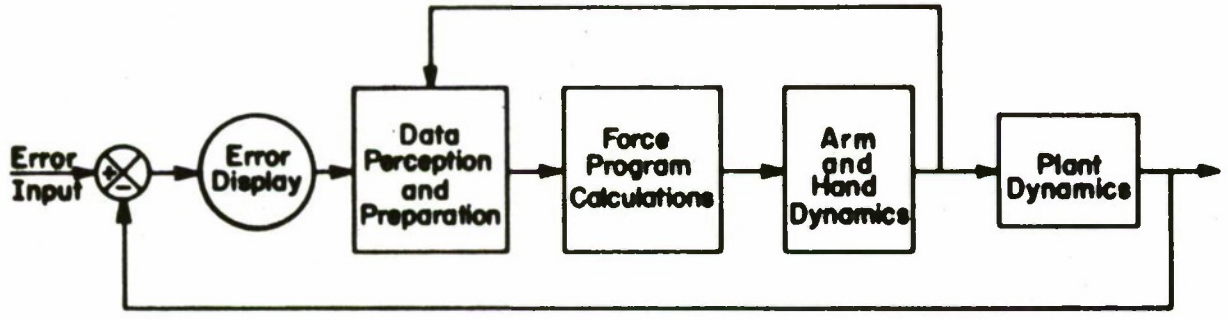


Figure 4. Adapted Model Form.

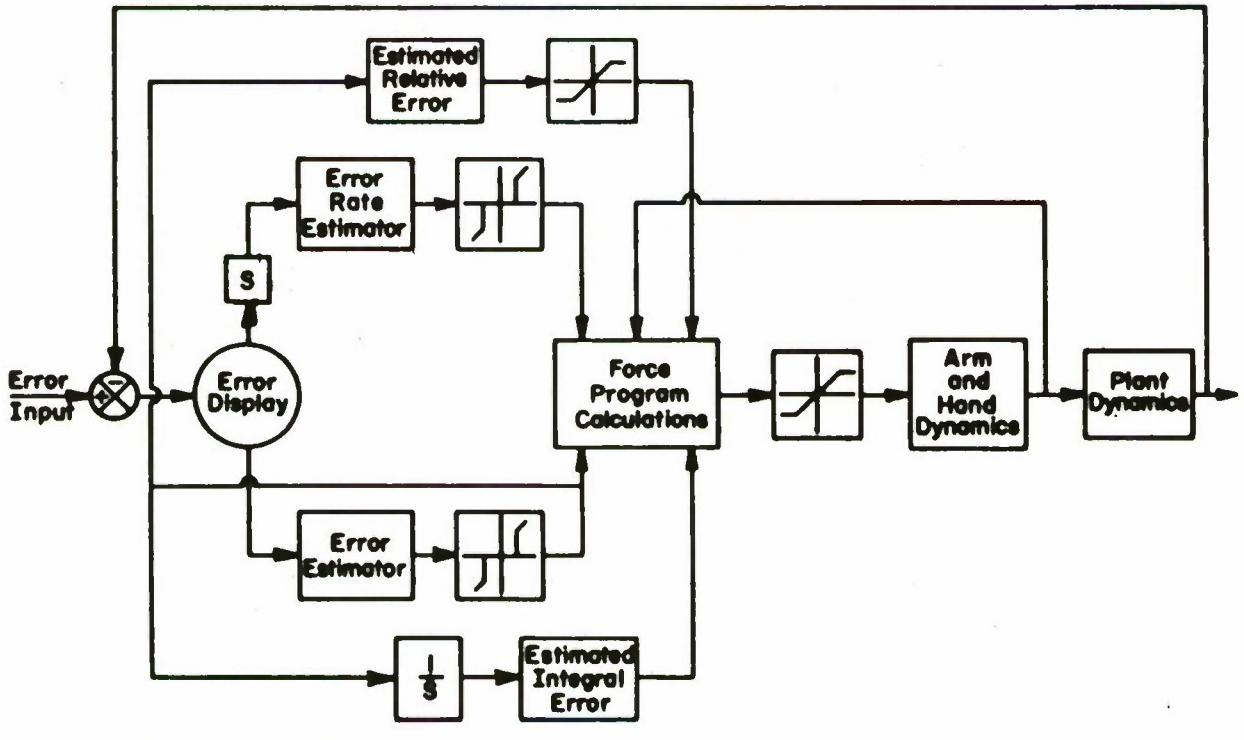


Figure 5. Data Perception and Preparation.

each mode of operation. The error and error rate estimator blocks represent functions described by

$$e^* = e + ne$$

and $$\dot{e}^* = \dot{e} + n\dot{e}$$

where e = true tracking error

$\dot{e}$ = true tracking error rate

n = normally distributed pseudo-random numbers.

The model was simulated on a digital computer; perception and control functions were performed at intervals which were integer multiples of the original data sampling rate.

Gain Calculation Algorithms

With the accumulation of the data inputs shown in Figure 5, calculation of the force program for a particular interval may begin. The adaptive component of each gain is a function which compares the tracking error in the present interval to the tracking error in the interval just past, that is:

$$\Delta F = \frac{G_1 e^*_k}{|e^*_{k-1}|}$$

where ΔF = adaptive gain correction factor

e^*_k = estimated tracking error for interval k

G_1 = a negative scaling constant

Since G_1 is always negative, the sign of e^*_k always correctly indicates the proper sign for the total force program gain with type 1 plant dynamics; the type 1 control strategy involves no control reversals. The adaptive nature of this function is inherent in the fact that an increasing error produces an increasing gain of opposite sign to offset it. This function taken alone, of course, would result in gains of unwarrented size in the regions where the tracking error passes through zero. It is therefore important to limit it with a saturation function and provide another gain component, the size and sign of which are directly dependent on the observed tracking error in the given interval.

$$F_1 = G_2 \, e^*_k \, .$$

where G_2 is a negative scaling constant and e^*_k is as defined above. Next, consideration is given to the tendency of the human operator

to reduce control stick motion after adjusting to an error drift in the positive or negative direction. The function used is:

$$F_2 = G_3 \, I_c$$

where I_c = the time integral of controller displacement

G_3 = a negative scaling constant.

This function, in effect, adds a changing bias term to the force program. The bias tends to generate a drift toward the center position of the controller and to accentuate the swing of the controller when the tracking error changes sign. The value of G_3 is relatively small and quite influential in governing control response.

Finally, a term is included which accounts for the requirement that the operator maintain a biasing force to keep the controller handle at a position off center. This term is peculiar to the spring restrained controller used and the adjustment of the springs during a particular tracking run. It is of the form:

$$F_3 = \frac{G_4}{G_5 + |D_c|}$$

where G_5 = a positive constant

D_c = controller displacement

G_4 = a constant with the same sign as D_c

This term is seen to have a diminishing effect as distance from the controller center position increases, which is in keeping with the nonlinear nature of the spring centering system on the controller used.

The total force program calculation algorithm for type 1 control situations has now been described. The result is the sum of the parts.

$$F_T = \Delta F + F_1 + F_2 + F_3$$

where F_T = total gain for the given axis in the given interval.

The total gain is, of course, limited to what the arm and hand neuromuscular system can provide; therefore, a saturation nonlinearity is built into the model at the force program output.

Control Strategies

Manual control through a type 1 plant is characterized by a lack of control reversals, whereas control through a type 2 plant requires a control reversal for each time the neutral position is crossed. The human operator's strategy appears to be to initially correct for an error as for type 1 control. Because of the greater lag of the system response with a type 2 plant, this results in relatively high error rates and a larger than normal tracking error. The next significant control motion of a trained operator is to wait for the error rate to change sign. Then a control reversal is effected and force programs applied until either the error rate changes sign again or the controller reaches a predetermined position on the opposite side of the center point.

During periods of zero estimated error rate, no significant control actions are accomplished. There are, however, in both types 1 and 2, indications that periods of the absence of meaningful error signals result in what appear to be dither signals applied to the controller. This phenomenon is most easily observed in type 1 control when the deviation from the center point is small. The human operator's objective appears to be to reduce the error magnitude until it is too small to yield meaningful estimates of error and error rate. He will then move the controller so as to produce a series of small, though observable errors, correcting for each as it occurs. But he will not allow the error indicator to remain stationary for longer than about 0.5 seconds if there is no controller motion.

The period over which a calculated force program is applied is a variable integer in the continuous error model. Because of the incremental nature of the operations, the period of force application may be varied in increments of time equal to the original sample point spacing. With the given error signals, a force application period of one interval was found to be satisfactory. This parameter could be expected to change if the nature of the error signal changed; it is substantially less than the force application interval reported by Fu and Gould in the discrete error case.

It is also possible to vary the reaction time delay period of the continuous error model, but the effective delay period was found to be zero. It may in fact be some value between 0 and 33.3 milliseconds, but no doubt closer to the former. This result is also very different from the measured results in the discrete error case. It suggests to the authors that a human operator operating with a continuous error signal is calculating a force program for use in the next interval while executing a precalculated program in the present interval.

The final step in the force program application is represented by the block labeled arm and hand dynamics in Figure 4. Analysis of physiological data by Fu and Knoop (6) and further work by Fu

and Gould (5) indicated that following transfer function was a reasonable approximation of the characteristics of the arm, hand and controller dynamics for the discrete error case.

$$\frac{X(s)}{F(s)} = \frac{1}{ms^2}$$

where $X(s)$ = distance moved

$F(s)$ = force applied to the controller

m = effective inertia of arm, hand and controller.

The present study indicates that a transfer function of the form

$$\frac{X(s)}{F(s)} = \frac{1}{ms}$$

is more nearly correct for the continuous error case. It was found that the extremely high gains required for operation with the former arm and hand dynamics resulted in chronic system instability, whereas the latter function allowed satisfactory operation. It is obvious that the difference in the controllers used could account for some of the change, and the functioning of the arm and hand neuromuscular system under essentially continuous rather than discrete movement could also affect the outcome. Unfortunately, it is impossible to conclude without further experiments which is the dominant reason for the change.

Adapted Model Response

Figures 6 and 7 show plots of the type 1 and type 2 model responses and the human operator responses to the same error signals. The reader should keep in mind that the human operator response to a given error signal is never exactly the same and, in fact, can appear quite different when compared to the results of earlier tests with the same error input. Point to point comparisons are therefore not very meaningful. General characteristics should be examined instead. A manual control system may be particularly sensitive to certain aspects of a human operator's control response, and this may lead one to look for certain characteristics such as total controller travel, number of controller reverses, maximum error, or some other measure. In this case, the use of a model rather than a human subject has an advantage in that the production of controller output forms of special interest may be more readily achieved by adjusting parameters of the model than by trying to give instructions to the human operator.

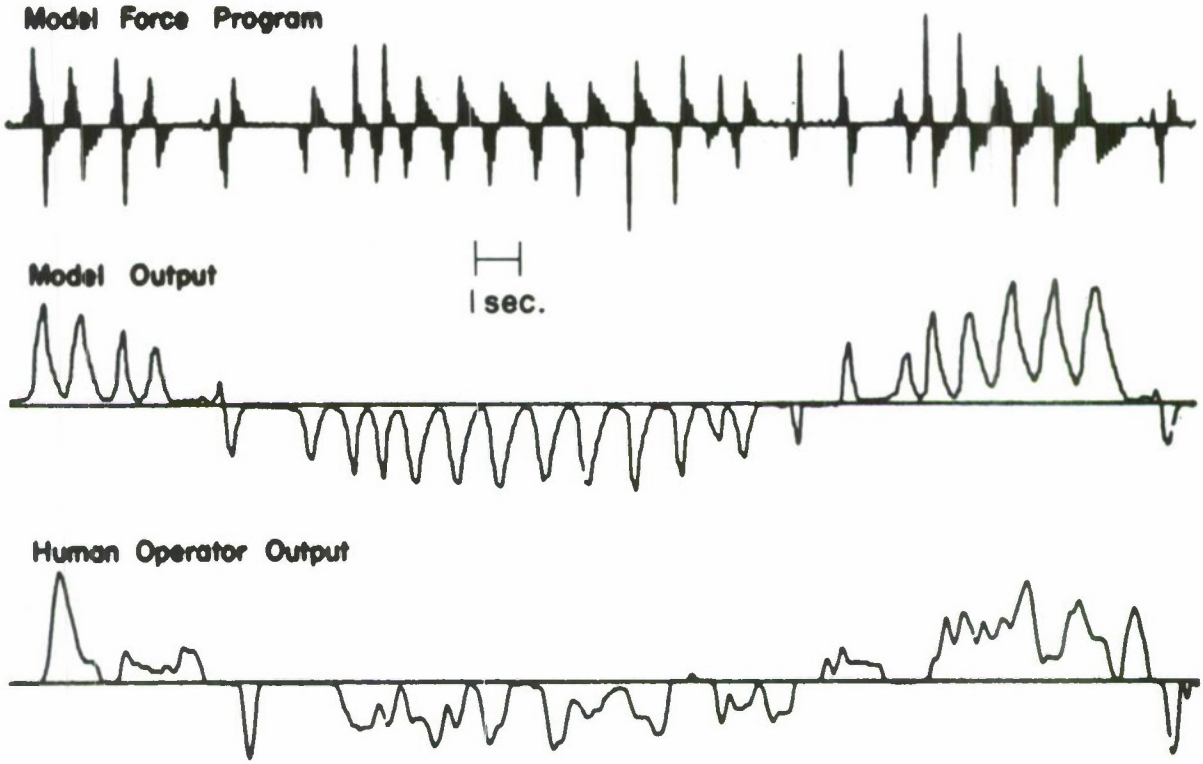


Figure 6. Type 1 Adapted Model Response.

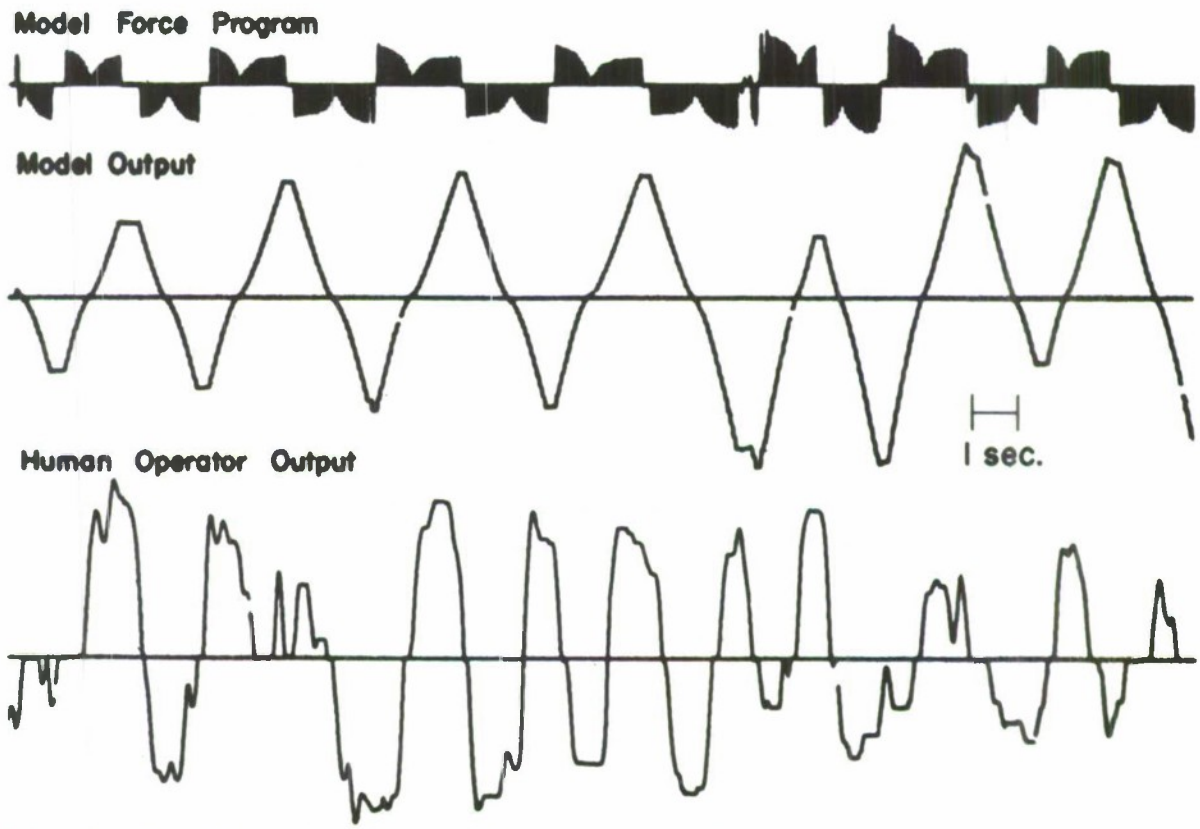


Figure 7. Type 2 Adapted Model Response.

Changes in Plant Dynamics

In the foregoing discussion, the word adaptive was used to describe the behavior of the single mode models with respect to changes in plant gain. This adaptation takes place within the framework of the control strategy provided in each of the models. The adaptive function to be considered now, however, concerns the changing of control strategies. Either of the two time invariant models developed in this study exhibits gain adaptability when active alone, and together with a pattern recognition device, as in Figure 8, they act to maintain system stability in the event of changes in either plant gain or form.

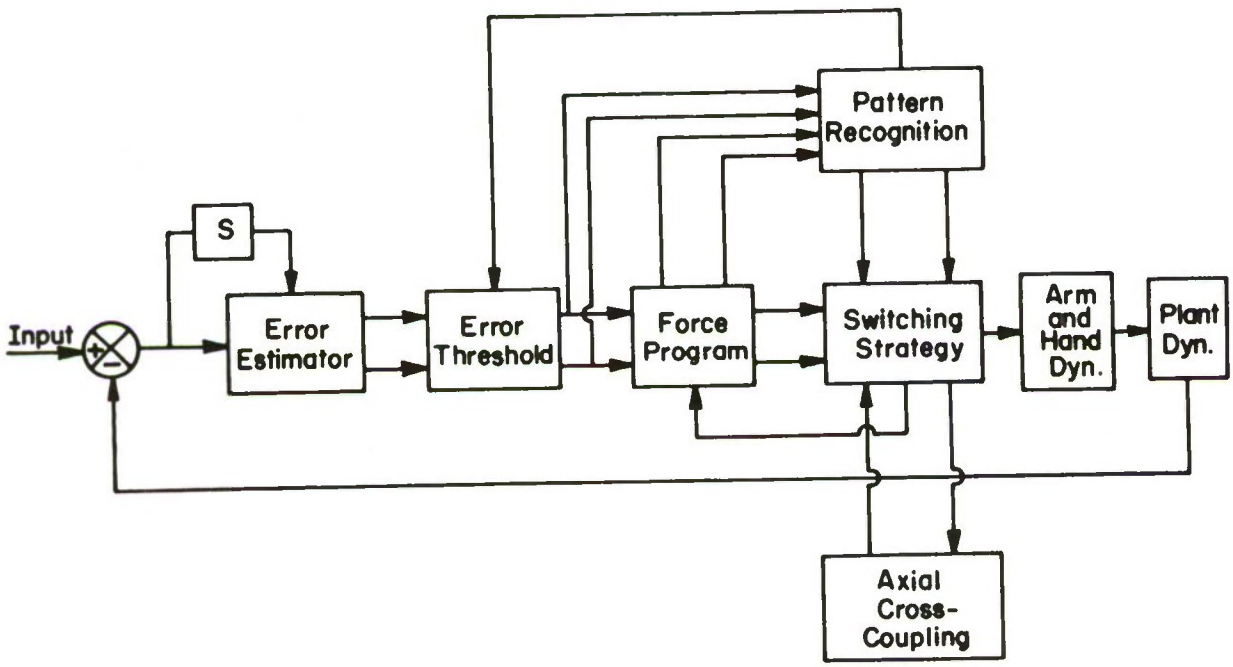


Figure 8. Adaptive Controller Model.

Changes in the form of the plant dynamics in a two dimensional system were allowed to occur both synchronously and asynchronously on the two axes. The former situation results in homogeneous dynamics at all times, while the latter involves the interesting problem of heterogeneous control strategies. In particular, the type 1 strategy, which results in the control output shown in Figure 6, is modified in some manner when it is exercised on only one axis, the other axis requiring type 2 strategy. The mechanism of the change is not clear, although it appears that the type 2 strategy is dominant. Examination of tracking records indicates that the controller motions typical of type 1 plants in the homogeneous case are still present in the heterogeneous case, but

the movements may be spaced farther apart. There appear to be more periods of zero or small gain, which correspond to similar periods in the type 2 controller movement records. Unfortunately, the behavior is not consistent in the same subject and varies widely between subjects, depending in part on the state of training. In spite of the apparent lack of characteristics amenable to analysis in type 1 heterogeneous tracking records, the authors obtained some interesting results by allowing the model parameter set obtained for type 2 tracking rules to be applied to the type 1 model. Figure 9 shows the results of this when the same error signal used for Figure 6 is applied. The differences between these two type 1 performances suggest to the authors that a portion of the cross coupling effect observed in heterogeneous tracking records is due to the fact that uniform standards for data acquisition (perception) errors and deadband limits prevail at a given time for both axes. The generally larger errors associated with type 2 control therefore govern control motions in the heterogeneous case.

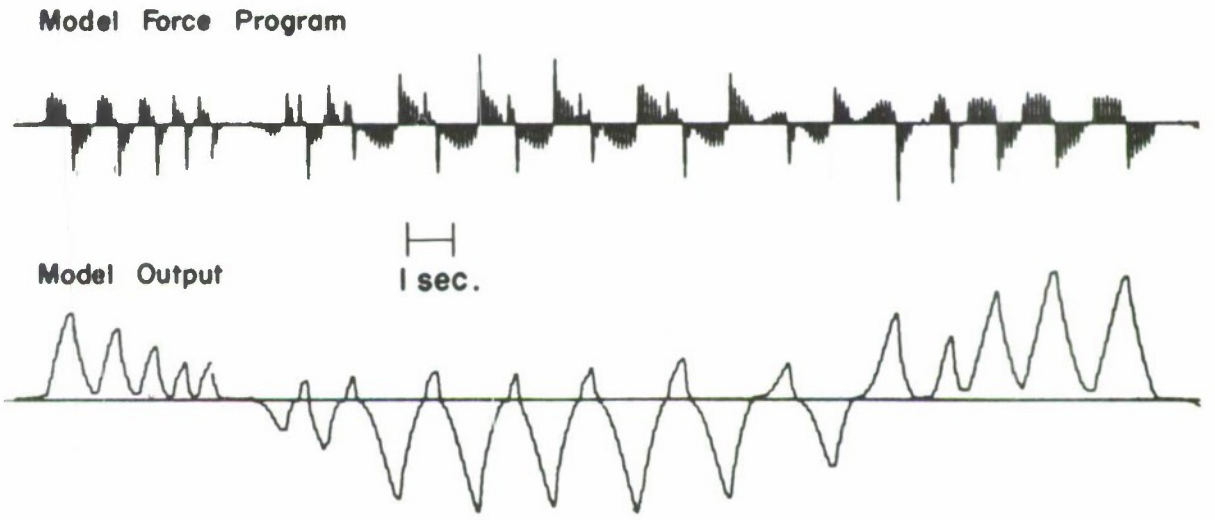


Figure 9. Modified Type 1 Model Performance.

Another possible cause of some observed behavior in the type 1 heterogeneous case is the rapid switching of control strategies. The pattern recognition device which models the perception of change process in this study will be discussed in greater detail in the next section. It is pertinant to note here though, that

albeit the performance of the machines tested was good, it was not perfect. Based on the test results, one could expect from 5% to 10% erroneous decisions in both type 1 and type 2 control. The effect of such a decision is to immediately change the model control strategy for the appropriate axis. The change can not remain in effect for long, in general, because either strategy exercised with the wrong plant dynamics results in system instability. But the occasional lapses might result in some of the apparently hybrid responses seen in both types of control.

The smoothest change in control strategy is observed in homogeneous tracking problems. An operator's total transition can be conveniently partitioned into the following four periods, which are described by Weir and Phatak (7).

Pretransition steady state

Retention of pretransition operator adaption

Optimal control

Adjustment of posttransition steady state.

Because of the structure of the model proposed herein, the above sequence occurs naturally, without any additional manipulations required. Figure 10 shows model and human operator responses to the same error signal.

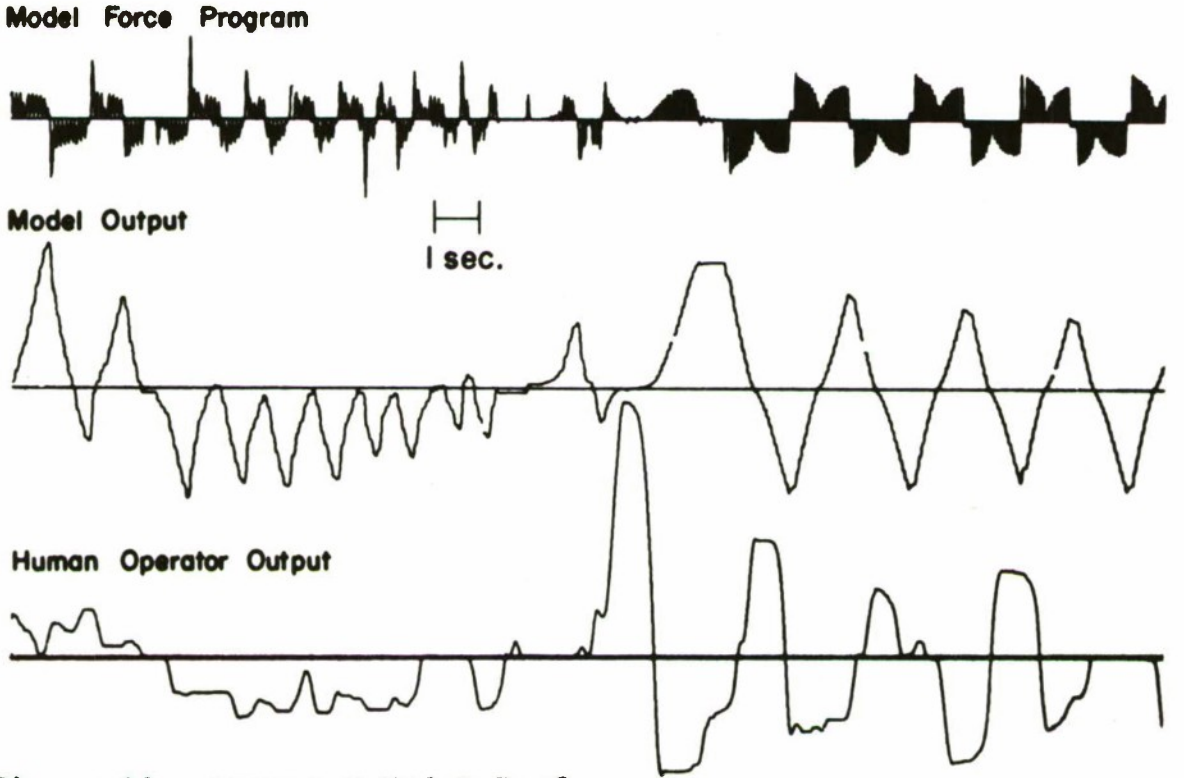


Figure 10. Type 1-2 Model Performance.

The model control strategy and plant dynamics were simultaneously driven to type 2 at the correct time through the use of a synchronizing pulse in the original data. One can observe in this interesting plot the gain adaptation responses of both types 1 and 2 strategies. Adaptation is required, of course, since model responses were calculated in blocks of 33 seconds. Since the error was continuously varying within a tracking period, consideration of an interior portion of a period would result, in general, in a step function error increase from zero to the value at the point of entry. The model response to this excitation reveals a tendency toward a time-optimal or bang-bang strategy during the transition activity, followed by a well behaved, steady state, type 1 response. A change in plant dynamics and control strategy yields another example of somewhat modified time-optimal adaptation to the new situation. The steepness of the rises and falls of the model output is limited by the gain which can be applied without causing system instability and the damping which is furnished by the spring centering force portion of the gain calculation algorithm.

The retention of either control strategy into a period where the use of it would result in system instability is allowed by the pattern recognition device, but as soon as the tendency toward instability is sensed, the strategy is switched. Phatak and Bekey (8) use a somewhat similar procedure, but with more than 2 modes. Figure 11 shows the form of the pattern recognition device used, and Figure 12 shows the results of training with a

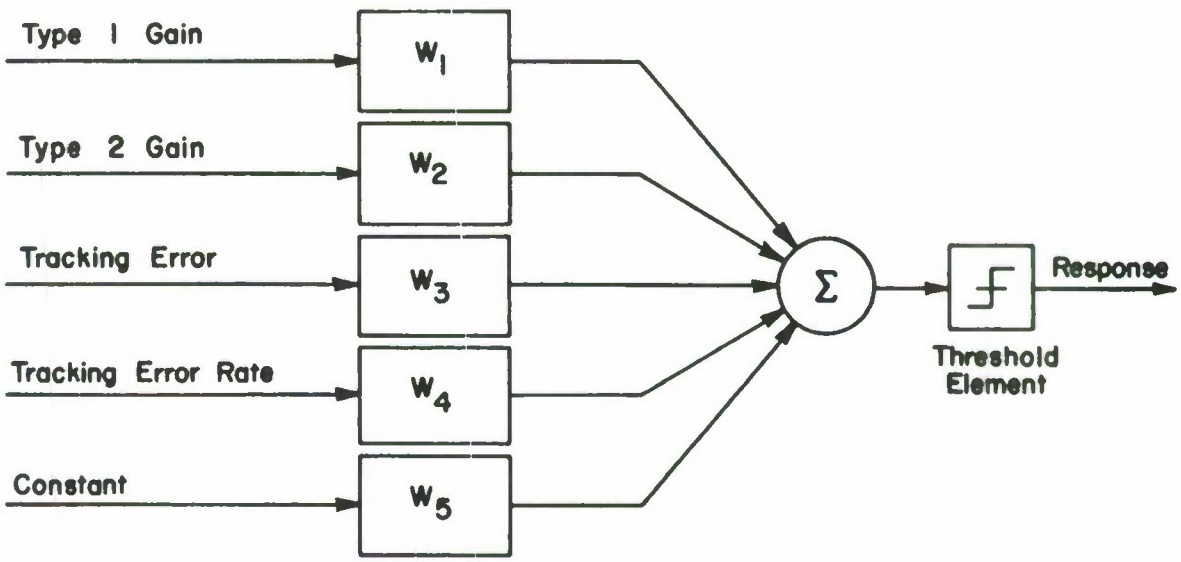


Figure 11. Linear Pattern Recognition Machine.

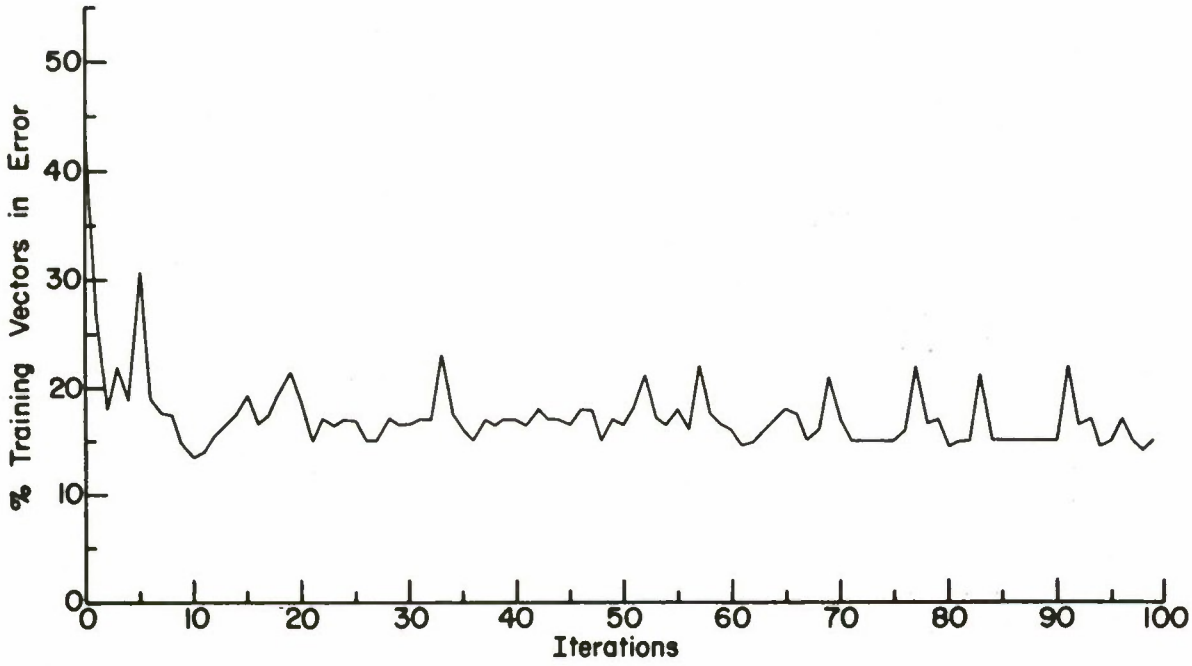


Figure 12. Convergence of Linear Machine.

set of 200 patterns generated by the model. The training procedure is described in (9). The performance of the machine suggests that the decision process should be quite rapid, and tracking tests show this to be true.

Concluding Remarks

A pattern recognizing, mode switching model of a human controller has been proposed. The performance of the model suggests that its basic form is adequate for the plant dynamics considered. Improvement within the framework proposed could be made in at least three areas.

1. Addition of more modes, so as to allow use of the model with a greater variety of plant dynamics.

2. Improvement of the pattern recognition algorithm to allow more complex partitioning of the pattern space.

3. Revision of the assumed transfer function (1/ms) for the arm and hand dynamics to more closely reflect the characteristics of muscle tissue under varying loads and excitations.

Bibliography

1. McRuer, D. T., Graham, D., Krendel, E. S. and Reisener, W., "Human Pilot Dynamics in Compensatory Systems," Air Force Flight Dynamics Lab. TR-65-15, Wright-Patterson AFB, Ohio, July 1965.

2. McRuer, D. T., Hofmann, L. G., Jex, H. R., Moore, G. P., Phatak, A. V., Weir, D. H. and Wolkovitch, J., "New Approaches to Human-Pilot/Vehicle Dynamic Analysis," Air Force Flight Dynamics Lab. TR-67-150, Wright-Patterson AFB, Ohio, Feb 1968.

3. Bekey, G. A., "The Human Operator as a Sampled Data System," IRE Trans on Human Factors in Electronics, HFE-3, pp. 43-51, Sept 1962.

4. 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, Jan 1965.

5. Fu, K. S. and Gould, E. E., "An Adaptive Pattern Recognizing Model of the Human Operator in a Time Varying Control Task," Purdue University TR-EE66-8, School of Electrical Engineering, Purdue University, Lafayette, Indiana, Jun 1966.

6. Fu, K. S. and Knoop, D. E., "An Adaptive Model of the Human Operator in a Control System," Purdue University TR-EE64-15, School of Electrical Engineering, Purdue University, Lafayette, Indiana, Sept 1964.

7. Weir, D. H. and Phatak, A. V., "Model of Human Operator Response to Step Transitions in Controlled Element Dynamics," NASA CR-671, Jan 1967.

8. Phatak, A. V. and Bekey, G. A., "Model of the Adaptive Behavior of the Human Operator in Response to a Sudden Change in the Control Situation," IEEE Trans. on Man-Machine Systems, MMS-10, No. 3, pp. 72-80, Sept 1969.

9. Nilsson, N. J. Learning Machines, McGraw-Hill Book Company, Inc., New York, 1965.

AN INPUT ADAPTIVE, PURSUIT TRACKING MODEL OF THE HUMAN OPERATOR *

John R. Ware
Simulation Research Center
The University of Michigan

A method is presented for modeling the human operator's input adaptive behavior. The main feature of the model is the simplicity of form and simulation either hybrid, analog, or digital.

In tracking random and pseudo-random continuous signals the human operator (H. O.) must attempt to predict the future state of the signal based on his observations up to the present. The prediction time is made up of the H. O.'s reaction time plus the time required for his control effort to cause the desired effect. In the simple model proposed it is assumed that the H. O. uses only position, $x(t)$, and velocity, $\dot{x}(t)$, information from the signal to make the best linear mean square prediction, $\hat{x}(t+\lambda)$, of $x(t+\lambda)$. If the signal is a Gaussian random process then the best linear mean square predictor is the best predictor for any error criteria that is symmetrical and a non-decreasing function for non-negative argument.

The problem then is: Choose α, β so that:

$$\hat{x}(t+\lambda) = \alpha\ x(t) + \beta\ \dot{x}(t)$$

minimizes:

$$E\left\{\left(x(t+\lambda) - \alpha\ x(t) - \beta\ \dot{x}(t)\right)^2\right\} = \overline{\epsilon}^2$$

If $x(t)$ has zero mean then the orthogonality principle can be used to show that:

$$E\left\{\left(x(t+\lambda) - \alpha\ x(t) - \beta\ \dot{x}(t)\right)x(t)\right\} = 0$$

$$E\left\{\left(x(t+\lambda) - \alpha\ x(t) - \beta\ \dot{x}(t)\right)\dot{x}(t)\right\} = 0$$

---

*This research was supported in part by NASA Contract No. NASr 54(06).

If x(t) is stationary then this becomes:

$$\alpha = R(\lambda) \,/\, R(0)$$

$$\beta = R'(\lambda) \,/\, R''(0)$$

$$R'(\tau) = \frac{d}{d\tau} R(\tau)$$

Because if x(t) exists, then $R'(0) = 0$. Suppose the signal is sampled periodically with sample period λ resulting in the sequence $x_0, x_1, \ldots$ where $x_i = x(i\lambda)$. $R(\lambda)$ is estimated by $\hat{R}(\lambda)$ where:

$$\hat{R}_N(\lambda) = \gamma \hat{R}_{N-1}(\lambda) + (1-\gamma) X_N X_{N-1} \qquad 0 \leq \gamma < 1$$

It is easy to show that this is an unbiased estimator.

A similar procedure is used to estimate $R(0)$, $R'(\lambda)$, $R''(0)$ i.e.

$$\hat{R}_N(0) = \gamma \hat{R}_{N-1}(0) + (1-\gamma) X_N^2$$
$$\hat{R}'_N(\lambda) = \gamma R'_{N-1}(\lambda) + (1-\gamma) X_N \dot{X}_{N-1}$$
$$\hat{R}''_N(0) = \gamma R''_{N-1}(0) + (1-\gamma) \dot{X}_N^2$$

From which

$$\alpha_N = \hat{R}_N(\lambda) \,/\, \hat{R}_N(0)$$
$$\beta_N = \hat{R}'_N(\lambda) \,/\, \hat{R}''_N(0)$$

It would be very difficult, if not impossible to determine if α_N, β_N were truly unbiased estimators of α, β but experimental evidence is strong that they are.

Since the past is weighted geometrically more recent observations affect α_N, β_N more strongly than past observations. Therefore if some characteristic of the input, such as cutoff frequency, is changed then α_N, β_N change also. Convergence rates depend on γ, the sampling rate, and the degree of change is α, β.

IMPLEMENTATION AND EXPERIMENTATION

In practice the input signal is sampled with a period of $\lambda/2 = E$. Mathematical operations and delay are performed on a digital computer and an output is generated every E seconds with constant velocity segments between samples.

In order to compare model performance with H.O. performance it is necessary to have an experiment in which the H.O. does practically nothing else but attempt

to track an input signal. For this reason control of a straight gain plant with a force stick is being used for the preliminary experiments. In this type of experiment a prediction interval of $.25 < \lambda < .3$ fits most subjects.

The figures show both model and H.O. tracking signals which undergo step transitions in characteristics.

AFTERWORD

Thus far the model has shown itself to mimic human tracking for a variety of continuous signals such as random process, sine waves plus noise, sine waves, and triangular waves, even in its simple form. A more complete version of the model could be made to handle inputs with non-zero mean, include a compensatory loop, include acceleration information, and be stochastic with a "remnant" dependent on input signal. Applications of the procedure for the modeling of human operators in control of plants with time delay is an obvious extension.

HARDWARE

The model is presently set up on a hybrid system consisting of an AD-4 analog computer tied to a PDP-9 digital computer. The model itself requires 2 (two) amplifiers and, not counting A to D and D to A conversions, only 17 arithmetic operations. It would be easy to do an all analog version of the predictor which would require at most 9 track store devices, 8 multipliers, 5 integrators and 2 summers. If one is willing to accept a Pade approximation to a delay it is possible to have a continuous version of the model with quite a saving in hardware. Included is a circuit which can be used to compute a typical function, $R(\lambda)$, required by the model.

REFERENCES

1. Bruce, M. M., "Estimation of Variance by a Recursive Equation," NASA Technical Note NASA-TN-D-5465, October, 1969.

2. Papoulis, A., "Probability, Random Variables, and Stochastic Processes," McGraw-Hill, 1965.

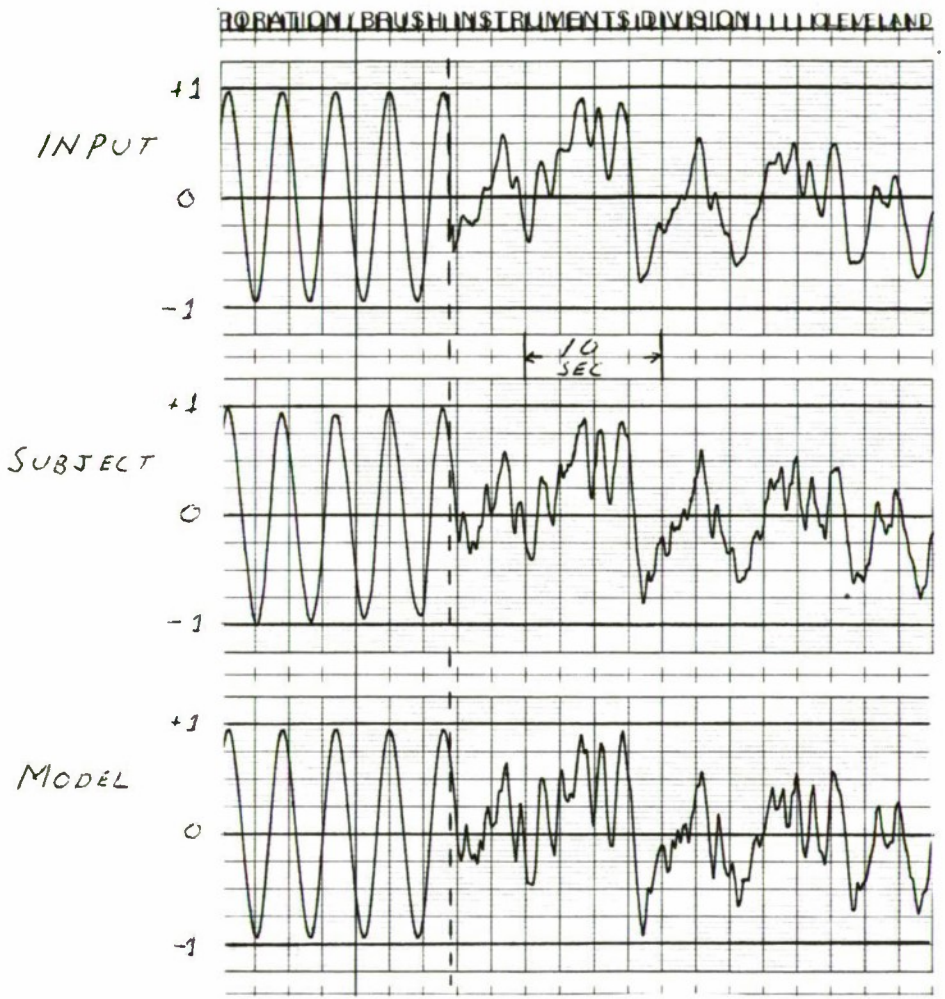


FIG. 1 TRANSITION FROM 1.5 RPS SINE WAVE TO 3.14 RPS CUTOFF NOISE.
$\gamma = .01$

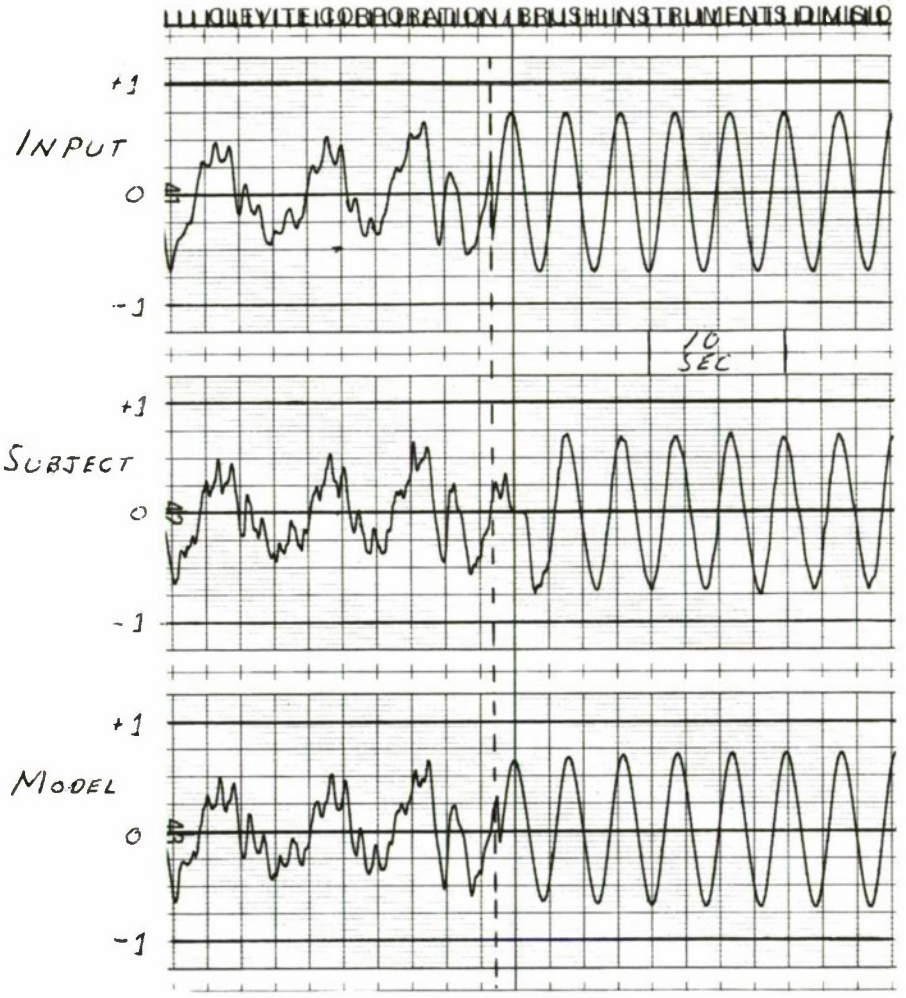


FIG. 2. TRANSITION FROM 3.14 RPS CUTOFF NOISE TO 1.5 RPS SINE WAVE.
$\gamma = .01$

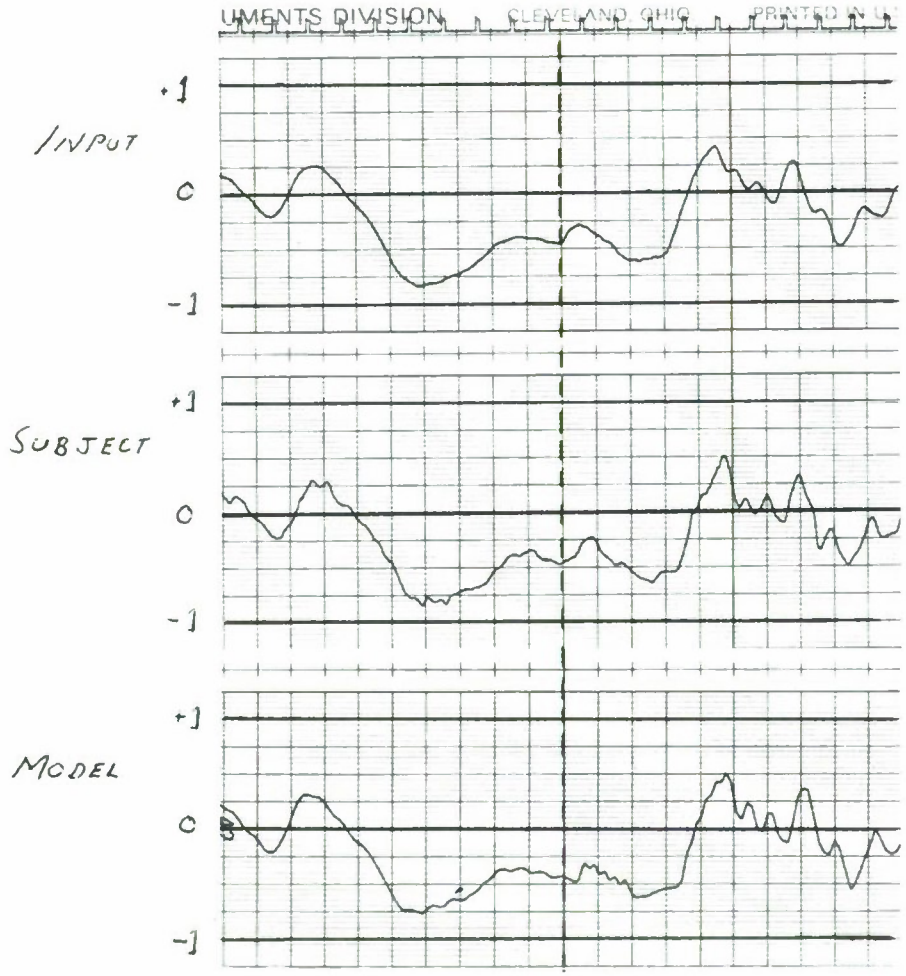
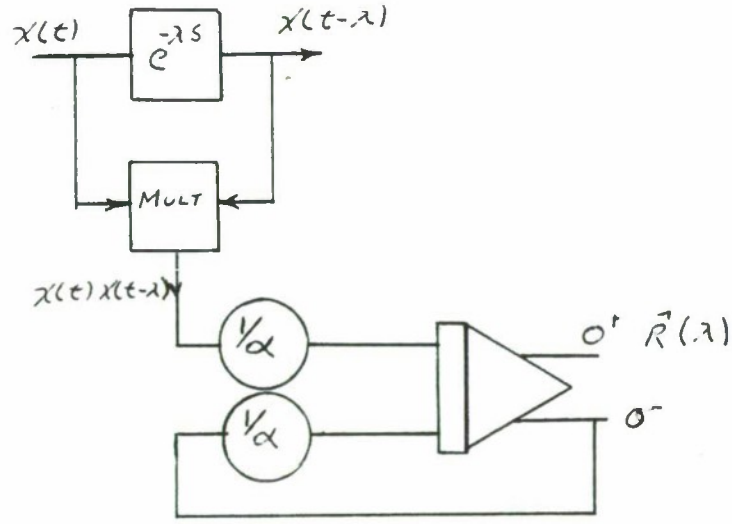


FIG. 3 TRANSITION FROM 1.57 TO 3.14 RPS CUTOFF RANDOM SIGNAL. $\gamma = .01$



$$E\{\hat{R}(\lambda)\} = R(\lambda)\left[1 - e^{-\frac{t}{\alpha}}\right]$$

FIG. 4. ANALOG COMPUTATION OF $R(\lambda)$.

A DISCRETE STOCHASTIC OPTIMAL CONTROL MODEL OF THE HUMAN OPERATOR

Major H. M. Paskin
AF Flight Dynamics Laboratory
Wright-Patterson AFB, Ohio

ABSTRACT

A discrete stochastic optimal control model of the human operator is developed for the single-loop compensatory/pursuit tracking situation. The model generates signals corresponding to those in the physical closed-loop tracking situation. There is one primary model parameter which is varied to match model-experimental normalized root-mean-squared tracking error at a bandwidth of ω_B=1.0 rad/sec for an input which approximates a rectangular spectra. With this parameter fixed, the model then predicts normalized tracking error and power spectra of control loop signals across a range of input bandwidths of 0.5 to 2.0 rad/sec. The model is applied to simple first and second order controlled elements in both compensatory and pursuit display situations.

A comparison between model and experimental normalized tracking error and power spectral density data confirms the model capability of matching and predicting operator performance with sufficient correlation to warrant its application as a tool in manual vehicular control system design.

INTRODUCTION

The general approach to modeling the human operator in closed-loop tracking tasks has been to match the experimentally obtained output response of the operator with a response generated by a model subjected to a similar input. The models are usually defined mathematically by some linear operation on the input stimulus and the output response match is normally made in the frequency domain. The portion of the operator's output not linearly correlated with the input is called "remnant". Most constituents of the model are chosen to represent some observed or assumed physical, physiological or psychological operator characteristic such as time delay, neuromuscular behavior, indifference thresholds, gains, lags, and leads. These model parameters are then varied to obtain a "best fit" to the experimental tracking data. The number of parameters normally determines how well the model output can be made to match the actual operator data, the better fits requiring more parameters. This approach has been quite successful in the single-loop compensatory tracking task and has evolved as the quasi-linear random input describing function model. Attempts have been made to apply this technique to the pursuit tracking situation (Ref 8, 10) but none have developed a successful model.

The object of this investigation is to develop a model of the operator-vehicle control loop which will predict operator performance, reproduce the essential control characteristics of the operator, be applicable over a wide range of forcing function bandwidths and controlled elements and apply to both compensatory and pursuit tracking situations. The approach taken is to write the tracking loop equations in state equation form and synthesize an optimal controller to minimize a quadratic cost function. The synthesis is accomplished using linear discrete stochastic optimal control theory and the resulting optimal control model is construed to represent the total operator-display-controlled element tracking situation.

THE TRACKING SITUATION

The physical situation to be modeled is the single closed-loop tracking situation described by the functional block diagram shown in Fig. 1.

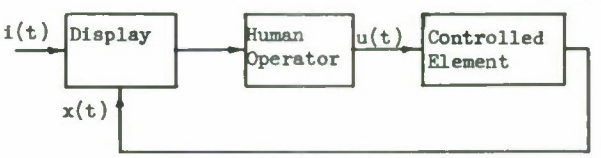


Fig. 1. Functional Block Diagram of Closed-Loop Tracking Task

No distinction will be made between pursuit and compensatory situations. This approach can be taken since there is no attempt made to relate the physical tracking situation to the model other than by defining the element to be controlled, the signal to bs tracked, and the quantity to be minimized.

The operator is presented a display (either pursuit or compensatory) and is given a manipulator with which he can control the output of a controlled elemsnt (vehicls) whose dynamics can be described by a linear constant coefficient differential equation of the form

$$\dot{X} = AX + Bu \qquad (1)$$

Eq (1) is a vector-differential equstion describing the response of the controlled element to a scalar input command u(t) generated by the human operator through the manipulator. The operator is generally instructed to control the plant output so as to minimize the mean-squared error between the displayed input and output signals (pursuit display) or to minimize the mean-squared value of the displayed error (compensatory display). The first stsp in developing the model is to establish the diffsrential equations rslsvant to the physical tracking situation.

Input Forcing Function. Tha forcing function chosen for this tracking task is a wide sense stationary stochastic process with a Gaussian amplitude distribution with zero mean whose power spectral density (PSD) is

$$\phi_{ii}(\omega) = \frac{K^4}{\omega^4+2\omega_B{}^2\omega^2+\omega_B{}^4} \tag{2}$$

Sample functions for the process described above can be obtained by operating linearly on a Gaussian amplitude white-noise process with two first-order filters, each of which is described by the following differential equation

$$\dot{x} + \omega_B x = Kw \tag{3}$$

where K = constant
ω_B = bandwidth (rad/sec)
w = sample function of whita noise process

Controlled Element. The controlled elemants considered in this investigation will be limited to plants governed by simpla first and second order linear differential equations of tha form

$$\dot{x} = ax + bu \tag{4}$$

and

$$\ddot{x} = bu \tag{5}$$

Manipulator. The manipulator, or control stick used in the experiments was an A.C. powered stiff-stick transducer. This manipulator was chosen because it has an essentially linear voltage output versus applied force and henca unwanted nonlinear control stick characteristics are eliminated. Thus the manipulator dynamics ara represented by a gain and no manipulator dynamics are included in the model.

THE DISCRETE OPTIMAL CONTROL MODEL

The next step is to combine all the differential equations which describa the operation of the control loop without the operator into one linear vactor differential equation of the form

$$\dot{X} = AX + Bu + W \tag{6}$$

where $\dot{X}$, X, B, and W are column vectors, A is an nxn matrix and u is the scalar control signal.

For illustrative purposes, the first-order controlled element will ba chosen, although the procedure is identical for highar order plants. Tha set of state equations is formad as follows

$$\dot{x}_1 = -\omega_B x_1 + Kw_1 \tag{7}$$

$$\dot{x}_2 = -\omega_B x_2 + Kx_1 \tag{8}$$

$$\dot{x}_3 = ax_3 + bu \tag{9}$$

where w is a sample function of a Gaussian amplitude white noise process, x_2 is tha desired forcing function input and x_3 is the controlled element output in response to tha control input u. Combine Eqs (7), (8), and (9) into vector form to get

$$\dot{X} = AX + Bu + W \tag{6}$$

where

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ 0 \\ b \end{bmatrix} \qquad W = \begin{bmatrix} Kw_1 \\ 0 \\ 0 \end{bmatrix}$$

$$A = \begin{bmatrix} -\omega_B & 0 & 0 \\ K & -\omega_B & 0 \\ 0 & 0 & a \end{bmatrix}$$

This completes the second step in tha model formulation. Nota that no spacial effort was made to formulate the state equations so the states represent some physical quantity observable or derivable by the human operator. The states are chosen as a logical consequence of writing differential equations in "normal form" (Ref 9:29). Now the observation equation will be defined and the plant equation (Eq(6)) will be discretized.

Define tha observation equation as

$$Y = X + \underline{V} \tag{10}$$

where Y, X, and $\underline{V}$ are column vectors. $\underline{V}$ is an "observation" vector whose components are linearly indapendent Gaussian amplitude white-noise processes. If the A matrix and B vector are constant, then Eqs (6) and (10) represent a time invariant linear systam and the complete solution for X(t) and Y(t) is given by (Ref 2:376)

$$X(t)=\Phi(t-\tau)X(\tau)+\int_{\tau}^{t}\Phi(t-\xi)[Bu(\xi)+w(\xi)]d\xi \tag{11}$$

$$Y(t) = X(t) + \underline{V}(t) \tag{12}$$

where t = observation time
τ = time u(t) was applied
$\Phi(t)$ = exp(AT) = state transition matrix

For computational purposes, it is desirable to develop a model which can be simulated on a digital computer. Consequently, Eqs (11) and (12) will be converted to a discretized form to which discrete linaar stochastic optimal control theory can be applied and which can then be simulated on a digital computer. Consider the situation where the control signal u(t) and tha plant noise vector W(t) are samplad every T seconds and applied to a zero-order hold network which maintains the value of u(t) and W(t) at the sampling instant for a time T. It is shown (Ref 8:126) that the response of a continuous time system to a sampled input of this form at the sampling instants (t=0,T,2T,...) is

$$X[(n+1)T] = \underline{A}X(nT) + \underline{B}u(nT) + \underline{W}(nT) \tag{13}$$

where $\underline{A} = \Phi(T)$

$$\underline{B} = \int_0^T \Phi(\xi)d\xi B$$

$$\underline{W}(nT) = \int_0^T \Phi(\xi)d\xi W(nT) \triangleq \underline{W}(n)$$

$$X(nT) \triangleq X(n)$$

A, B, and W(n) are computad using Eq (13) (Ref 7).

Remnant. Attempts to model the human operator by matching experimental data with a "linear" model have shown that there is always some portion of the operator's output which is not linaarly correlated with the input stimulus. This portion of the oparator output has been popularly called "remnant" and has been represented (Ref 4:119) as an additive white-noise process at the operator's output or as an observation noise (Ref 3) at the operator's input. The concept of remnant as an observation noise has been accommodatad by the obsarvation noise vector in Eq (10). To include the possibility of remnant as an additive noise to the control output, a Gaussian amplitude white-noise process w_2 ("motor" noise) will ba àddad to the control output u of the model. Thus, the model output is redefined to be

$$u'(t) = u(t) + w_2(t)$$

and Eq (13) is modified accordingly.

These observation and motor noises have been added to the model to investigata thair effect on model performance and as parameters which can ba varied to obtain correlation between model and human opertor performance. There will be no attempt made to associate the effects of thesa parameters with human physiological behavior.

Optimal Control Model. The plant and observation equations can be represented schematically as in Fig. 2. Note that a box labeled

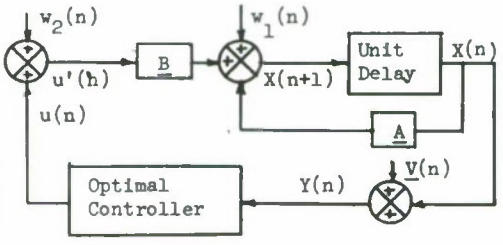


Fig. 2. Schematic Diagram of Optimal Control Model

"optimal controller" is included in this figure. The optimal controller is the mathematical construct from which the control u(n) is derived on the basis of the observed output Y(n). Thus this figure is a schematic representation of the proposed discrete stochastic optimal control model of the human operator in a closad-loop tracking task. The form of the optimal controller is now described.

The theory of discrata stochastic optimal control has been developed and documented. The specific case where the plant and obsarvation equations ara linear and the cost function is quadratic is derived in detail by Meier (Ref 5:19). To apply this theory to the control situation described by Eq (13), a cost function of the form

$$J = E[\sum_{n=0}^{N} X(n)^T V X(n) + u(n)^T S u(n)] \qquad (14)$$

must be stipulated. Since the control u for the single-loop tracking task is a scalar, the control weighting matrix S in Eq (14) is just a scalar and the matrix V in Eq (14) is formed such that

$$X(n)^T V X(n) = (\text{System Input-System Output})^2$$

or, for the example under consideration

$$V = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix} \qquad (15)$$

With suitable assumptions on tha mean and independence of tha observation and motor noise vectors (Ref 5:19) recursion formulas can be derived to generate the conditional mean of X(n) and the optimal control u(n).

$$u(n) = -G(n)\bar{X}(n) \qquad (16)$$

$$\bar{X}(n) = E[X(n) | Y^n, u^{n-1}] \qquad (17)$$

where Y^n represents all past observations and u^{n-1} represents all past controls. Thase recursion formulas can ba programmed to simulate the closed loop tracking task and as such represent the discrete stochastic optimal control model of the human operator.

DIGITAL SIMULATION

The recursion formulas (Ref 5:24) are mechanized on an IBM 7094 digital computer using Fortran IV language. Tha computer program generates N values of i(n), x(n), e(n), and u(n) where

i(n)=digitized value of system input at t=nT
x(n)=digitized value of system output at t=nT
e(n)=i(n) - x(n)
u(n)=digitized value of control input at t-nT

Tha power spectral densities (PSD) of these signals are generatad as a function of frequency and the normalized tracking error (NTE) defined as

$$NTE = \frac{\text{Variance of Error Signal}}{\text{Variance of Input Signal}} = \frac{\sigma^2 E}{\sigma^2 I}$$

is computed.

System-Model Parameters. The parameters in the model are tabulated below. These parameters are classed as either system or model parameters. The system parameters are thosa which are specified by the dynamics of the tracking task to be modeled. Tha model parameters are those which can be varied to change the model characteristics.

System Parametars

a = plant time constant
ω_B = input bandwidth
σ_I^2 = variance of input forcing function
b = plant gain

Model Parameters

N = number of samples
T = sampling period
S = control weighting in cost function
$\sigma^2 q$ = motor noise variance
$\sigma^2 v_i$ = observation noise variances
i=1,2,..,n.

The value of N chosen for model simulation is chosen as 4000 and the value of T is chosen to be 0.1 seconds to obtain PSD plots up to $\omega_{max}=\pi/T=10\pi$ rad/sec (Ref 1). Small observation noise variances are chosen to minimize their effects on the model since it was found that large variances increased NTE scores but had little effect on the gross characteristics of the model PSD data. The addition of motor noise increased NTE scores also, but altered the characteristics of the control spectra. It was found from the experimental data, though, that by varying the value of S, a satisfactory correlation between experimental and model data could be achieved without using motor noise variance as an active parameter. Hence the motor noise variance is also chosen to minimize its effect on the model output. Thus the primary model parameter to be varied is the control weighting S, all other model parameters being chosen as specified above.

EXPERIMENTS

The experiments conducted during this investigation were designed to validate the proposed model. To "validate" the model is to show that the NTE and PSD data obtained experimentally agree with the same quantities predicted by the model. Four trained subjects were used to obtain compensatory and pursuit tracking data for the first and second order plants considered over a range of input bandwidths of 0.5 to 20 rad/sec.

EXPERIMENTAL RESULTS AND DISCUSSION

For all controlled elements considered, a forcing function bandwidth of ω_B=1.0 rad/sec is chosen as the data point at which to match the model to the data. This value of ω_B was chosen on the basis of an analysis of the tracking data presented by McRuer. (Ref 4). The analysis reveals that because of psychophysiological limitations, the human operator may not operate in a linear manner across the range of forcing function bandwidths considered. "Indifference threshhold" effects occur at low frequencies and neuromuscular effects occur at high bandwidth frequencies. Thus a value of ω_B=1.0 rad/sec is chosen to establish a data point in the linear range of operator behavior.

Normalized Tracking Error. The results of the normalized tracking error experiments are shown in Figs. 3 and 4. These figures show the comparison between pursuit and compensatory tracking performance (as measured by NTE) for each of the controlled elements considered. These results agree with those obtained by Wasicko (Ref 10:44) and Reid (Ref 8:24), who conclude that in general, the NTE scores improve with pursuit displays for the more difficult controlled elements and the higher bandwidth inputs. These figures also compare the predicted and measured values of NTE based on a data match at ω_B= 1.0 rad/sec. In the K/S-1 and K/S^2 cases where the compensatory performance differed significantly from the pursuit, a model match is accomplished separately for each display condition. For the second-order controlled element (Fig. 3) the experimental NTE becomes greater than predicted for the compensatory case as ω_B increases. This trend towards higher than predicted NTE is predictable on the basis of compensatory tracking data collected by McRuer (Ref 4). McRuer found that operator remnant increases with increasing ω_B and hence operator output departs from its linear representation which in turn causes larger than predicted values of NTE. For the pursuit situation the operator NTE is lower than predicted at ω_B=2.0 rad/sec. This behavior might be attributed to a decrease in effective time delay as noted by McRuer (Ref 4) and attributed to the neuromuscular system. Fig. 4 shows the NTE results for the first-order controlled elements. The experimental mean values for these cases generally conform to the predicted values. At most of the data points the $\pm 1\sigma$ range of experimental NTE include the predicted values.

Thus although there are departures of the actual from predicted values in Figs. 3 and 4, they all have plausible explanations and consequently the predictions are close enough to warrant their use in evaluating the effects of input bandwidth changes on a given controlled element-display configuration.

Power Spectral Densities. Experimental and model power spectral densities were computed for the system input, output, error, and control signals and normalized spectra for system error and control were plotted. All spectra are normalized by $\phi_{ii}(0.1 \text{ rad/sec})$ as an approximation to $\phi_{ii}(0)$. Figs. 5 through 8 show some of the error and control spectra for the pursuit situation. In every instance the PSD is consistently above the model PSD at the higher frequencies (> 1.0 rad/sec). This is to be expected, as an analysis of the control PSD shows that the operator consistently has more control power than the model across the frequency spectrum, particularly in the 2-10 rad/sec frequency range. The general excess of control power can be attributed to an additive white-noise remnant (Ref 4, 8) and the emphasis at higher frequencies to neuromuscular system behavior (Ref 4). As would be expected, the characteristics of the error spectra do not vary with controlled element or bandwidth, but the control spectral characteristics vary in a manner which can be predicted by power spectral analysis and linear system theory. (Ref 6:347).

Figures 9 through 12 show some results of the compensatory tracking experiments. Note that the characteristics of both the control and error spectra are the same as for the pursuit case and conform to the predicted results. This is to be expected since regardless of the display situation, the task of the controller is to generate an input into the controlled element such that its output follows the system input. This task is described mathematically in the model by the cost function J and it may be inferred that the operator is trying to minimize

some subjective cost function in a similar manner. This minimization is independent of the type of display or controlled element. The form of the cost function in the model is constant as a function of type display or controlled element. Only the control weighting S changes to compensate for these changes.

CONCLUSIONS

As a result of the normalized tracking error and power spectral density data collected and analyzed, it is concluded that the model is valid over the range of controlled elements, forcing function bandwidths, and types of displays investigated. The model is easy to apply since the system equations can easily be derived once the type of input, controlled element, and manipulator are specified. These equations can then be used in the digital simulation of the optimal controller to generate system power spectra and normalized tracking error. The model-operator match is accomplished by varying the control weighting S to match the model NTE to measured or estimated operator performance for the type of display situation. System parameters can then be varied to predict performance in terms of NTE and PSD plots.

This system approach to modeling has provided new insight into the pursuit/compensatory tracking situations and has tended to unify the theory that the human operator behaves in some optimal manner while performing the single closed-loop tracking task. Furthermore, within the framework of discrete stochastic control theory the general form used for the single-loop control tasks investigated here might be extended to multiple axis/loop tracking. There is considerable data available against which this model extension could be compared. It is also

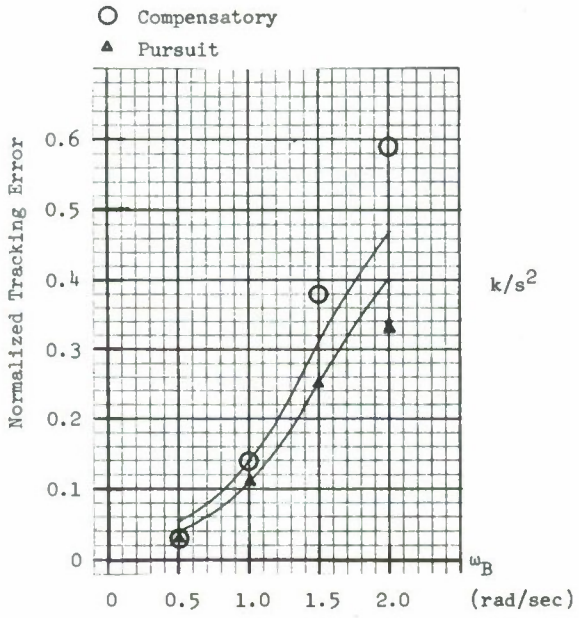


Fig. 3. NTE vs. ω_B Second Order Controlled Element

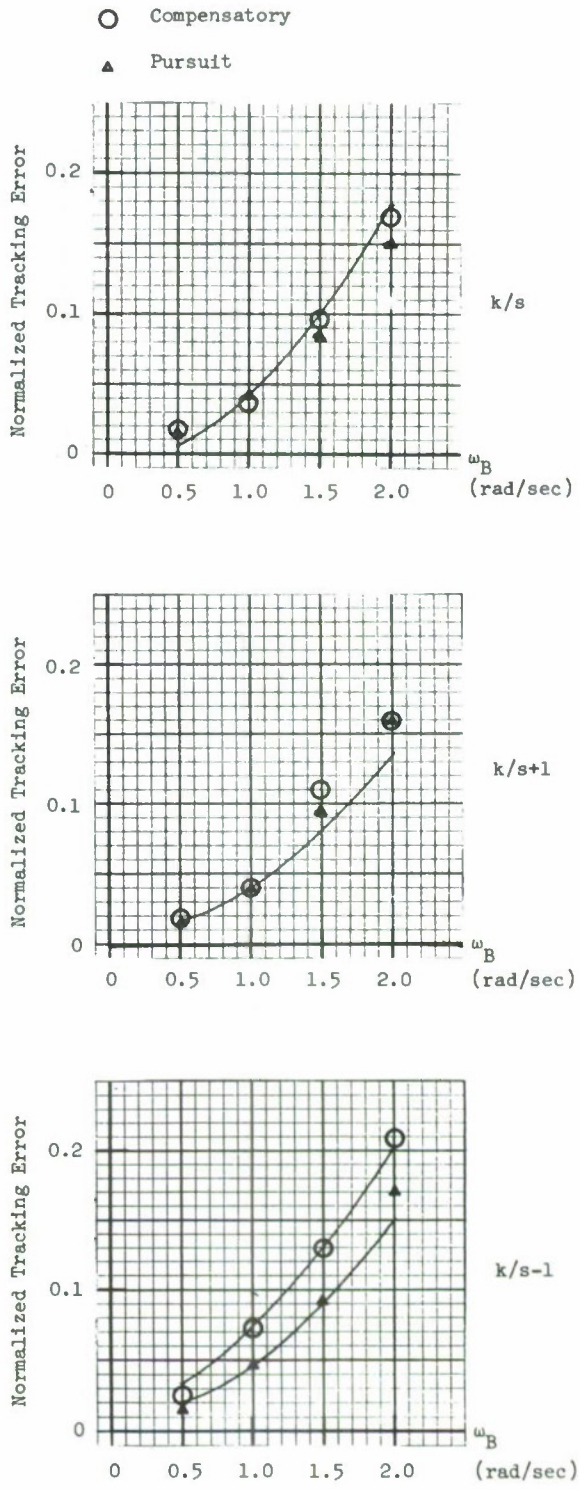


Fig. 4. NTE vs. ω_B First Order Controlled Elements

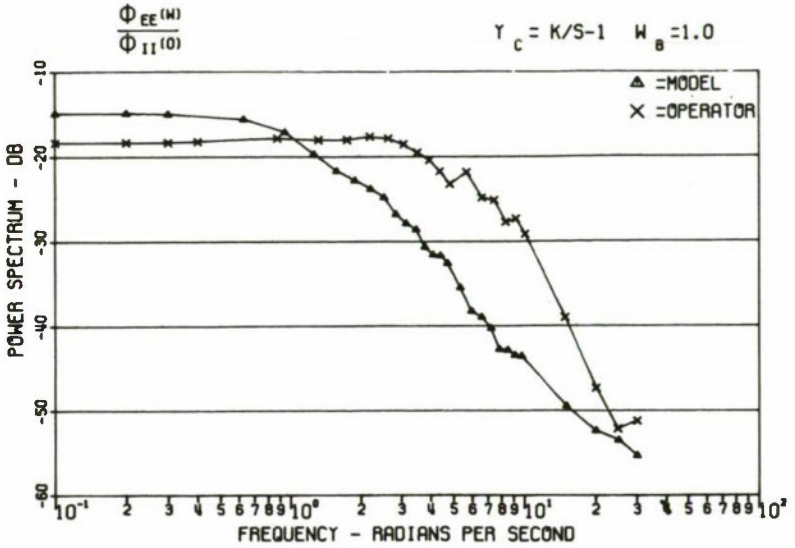


Fig. 5. Error Power Spectrum Pursuit

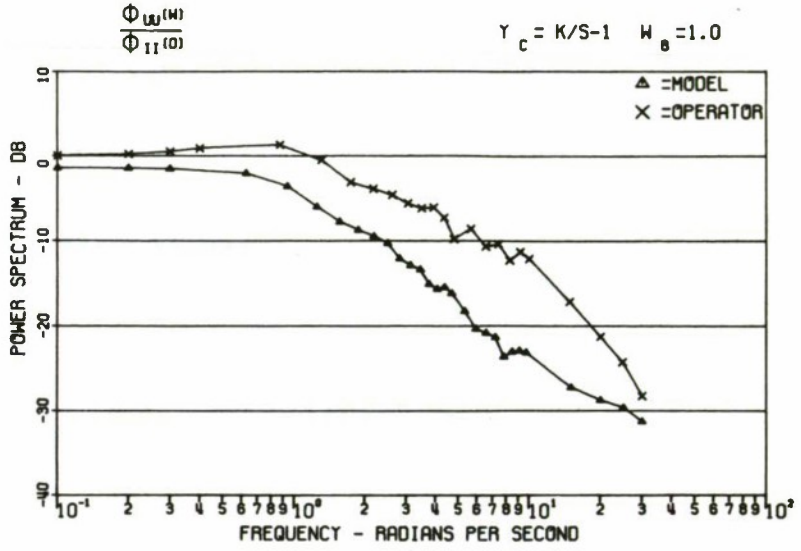


Fig. 6. Control Power Spectrum Pursuit

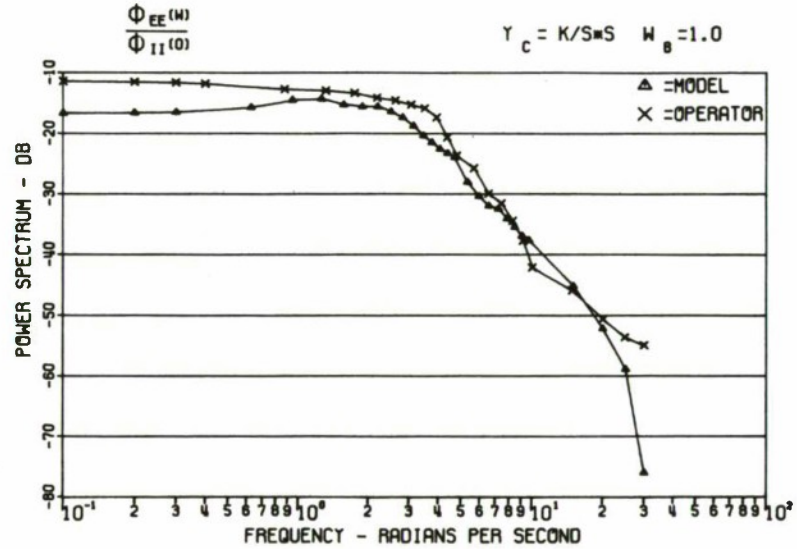


Fig. 7. Error Power Spectrum Pursuit

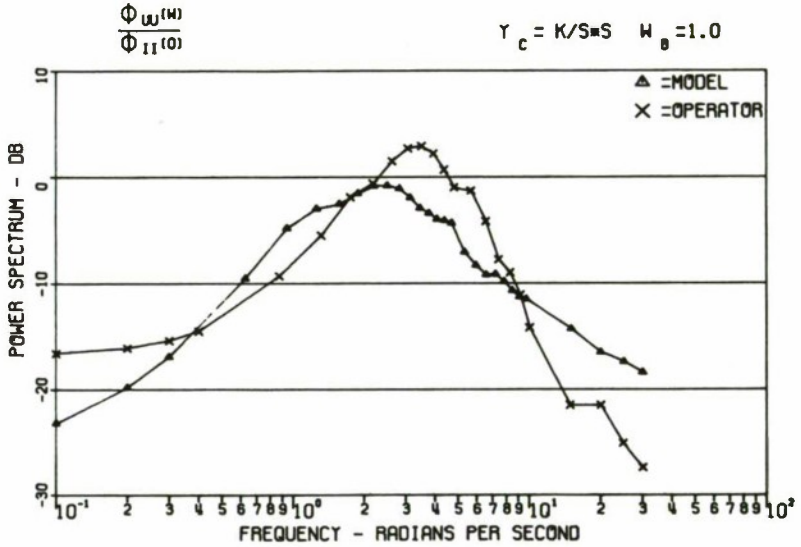


Fig. 8. Control Power Spectrum Pursuit

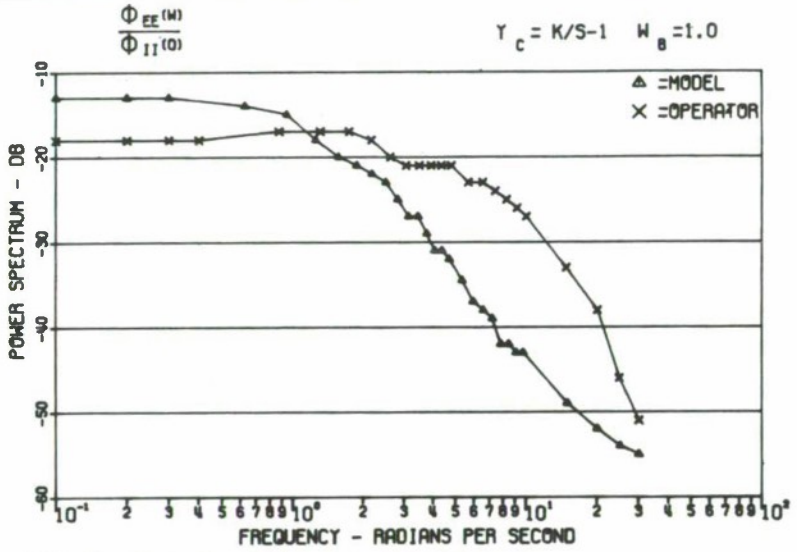


Fig. 9. Error Power Spectrum Compensatory

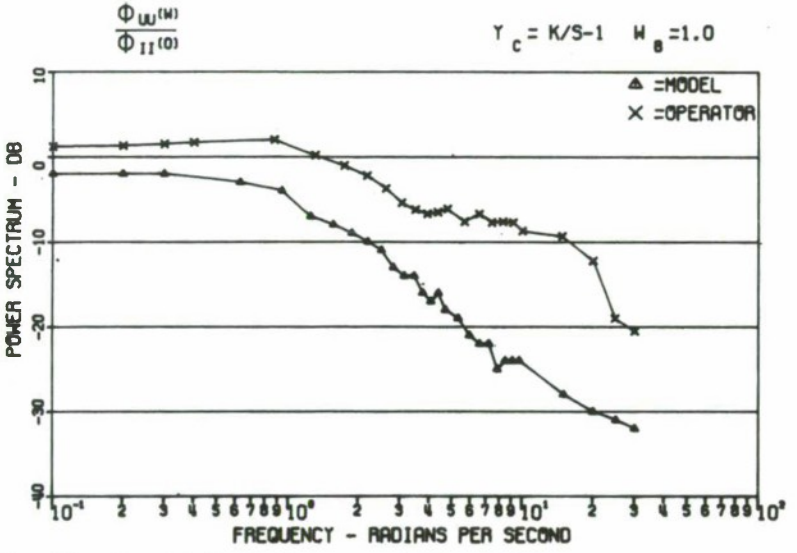


Fig. 10. Control Power Spectrum Compensatory

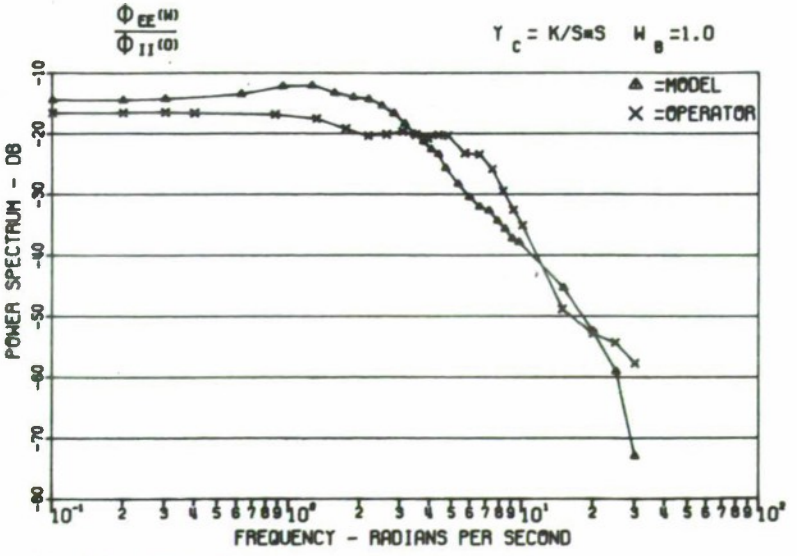


Fig. 11. Error Power Spectrum Compensatory

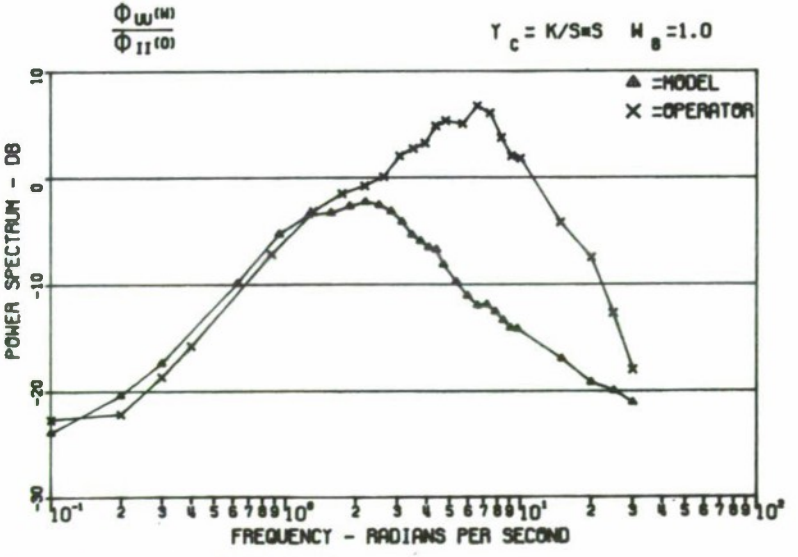


Fig. 12. Control Power Spectrum Compensatory

possible to apply the theory to the prediction of optimum sampling strategies for the multiple axis situation and this is another area to which the model might be extended.

REFERENCES

1. Bendat, J. S., and A. G. Piersol. Measurement and Analysis of Random Data. New York: John Wiley & Sons, 1966.

2. DeRusso, P. M., et.al. State Variables for Engineers. New York: John Wiley & Sons, 1966.

3. Levison, W. H., et.al. "A Model for Human Controller Remnant." IEEE Transactions on Man-Machine Systems, MMS-10: December 1969.

4. McRuer, D. T., et.al. Human Pilot Dynamics in Compensatory Systems. AFFDL-TR-65-15. Wright-Patterson AFB, Ohio: Air Force Flight Dynamics Laboratory, July 1965.

5. Meier, L. Combined Optimum Control and Estimation Theory. NASA Contractor Report 426. Washington: National Aeronautics and Space Administration, April 1966.

6. Papoulis, A. Probability, Random Variables, and Stochastic Processes. New York: McGraw-Hill Book Co., Inc., 1965.

7. Paskin, H. M. A Discrete Stochastic Optimal Control Model of the Human Operator, Ph.D. Thesis, Air Force Institute of Technology (To be published).

8. Reid, L. D. The Measurement of Human Pilot Dynamics in a Pursuit-Plus-Disturbance Tracking Task. UTIAS Rpt 138. Toronto: University of Toronto Institute for Aerospace Studies, April 1969.

9. Schwartz, R. J. and B. Friedland. Linear Systems. New York: McGraw-Hill Book Co., Inc., 1965.

10. Wasicko, R. J., et.al. Human Pilot Dynamic Response in Single-Loop Systems with Compensatory and Pursuit Displays. AFFDL-TR-66-137. Wright Patterson AFB, Ohio: Air Force Flight Dynamics Laboratory, December 1966.

A MODEL FOR TASK INTERFERENCE

By

William H. Levison

Bolt Beranek and Newman Inc.
Cambridge, Massachusetts

ABSTRACT

A model is presented for interference among multiple continuous manual control tasks. This model is based upon the assumption that multiple tasks are performed in parallel and that the human must share a fixed amount of central-processing capacity among the tasks. The equivalent observation noise associated with a subtask is shown to be related to the fraction of the controller's capacity allocated to that task. The model is able to predict with great accuracy the total performance scores measured in a set of multivariable tracking experiments. The model structure also accounts for the effects of multivariable tracking on the controller's describing function and remnant.

INTRODUCTION

Because the human can exert only a limited amount of physical or mental effort at any given time, his performance on a given psychomotor task degrades as he is required to perform more and more tasks simultaneously. Thus, multiple tasks tend to "interfere" with one another. Interference which results from limitations on the human's central-processing capability has been most commonly attributed to single-channel behavior; i.e., there has been postulated a single-channel somewhere in the central processor which must selectively "attend" to the various sensory inputs (Ref.1). Parallel-channel behavior, however, has also been considered in the literature; a multiband filter theory, for example, has been advanced to account for the behavior of the human observer when detecting

multiple auditory signals (Ref.2). The approach taken in this paper is to consider the human controller, as well as the human observer, as a parallel-channel processor of information.

This paper describes a model for predicting the interference that occurs when two or more continuous manual control tasks are performed concurrently. Only central-processing sources of interference are considered; performance degradation that is associated with visual scanning or intermittent control activity are not considered. The model development and validation presented in this paper are the principal results of a study on multivariable manual control performance recently completed for the NASA-Ames Research Center. Since considerations of space limit the author to presenting only the highlights of the experimental results, the reader is referred to Reference 3 for a fuller description of the validation experiments.

BACKGROUND

The control behavior of the human controller may often be represented adequately by a quasi-linear model which consists of two components: a describing function to account for the linear portion of the controller's response, and a remnant term to account for the remainder of the controller's output. Quasi-linear models for the controller have been well documented in the literature (Refs.4-8). Since the model for task interference is basically an extension of the model for controller remnant described in References 5 and 6, the remnant model is discussed briefly.

Remnant is modelled as an equivalent observation noise process; i.e., each sensory input variable used by the controller is considered to be disturbed by a white noise process. Since the human

controller is presumed to obtain both position and rate information from a single display indicator, the observation noise must be considered as a vector noise process. Remnant data obtained from a variety of single-axis control situations can be accounted for by an observation noise process of the form

$$\underline{R} = P \cdot (\underline{\sigma}^2 + \underline{\sigma}_T^2) \qquad (1)$$

where $\underline{R}$ is the injected white noise vector; P is the power density level of a normative white noise process, or "noise ratio"; $\underline{\sigma}^2$ is a vector composed of the variances of the displayed signals; and $\underline{\sigma}_T^2$ is the vector variance of noise processes internal to the controller. The latter term is necessary to account for the irreducible, or threshold, component of observation noise; this factor is particularly important when the display is viewed peripherally.

For the most part, we have found the noise ratio P to be remarkably consistent across experimental conditions. The near invariance of this measurement suggests that remnant – at least under the idealized conditions that have been investigated in the laboratory – arises largely from a central-processing type of disturbance common to all tracking tasks (such as time-variational disturbances of controller gain or time delay). The potential correspondence of equivalent observation noise to central-processing noise forms the basis for our model of task interference.

The model for interference developed in this paper has been incorporated into the optimal-control model for human performance developed by Baron and Kleinman (Refs.7 and 8) in order to provide a means by which the effects of task interference on system performance may be predicted. The reader is referred to the above references for a detailed description of the optimal control model. This model has been used to verify the task interference model.

MODEL DEVELOPMENT

Basic Assumptions

The model is founded on the following primary assumptions: (a) Multiple tasks are performed in parallel, not in sequence. (b) The controller has a relatively large fixed number of "information-processing channels" to distribute among his various tasks. (c) Each of these channels is perturbed by a white, Gaussian noise process which is linearly uncorrelated with all other noise processes and with system variables. The noise levels are proportional to signal variance.

The assumption of parallel processing is not crucial to the arguments presented in this paper. We could as well assume a very rapid internal scanning mechanism, since, in the limit of arbitrarily rapid scanning, the sequential and parallel models appear to lead to the same results. The assumption of parallel processing has been adopted primarily for mathematical convenience (and also because time-domain and frequency-domain analysis of the author's manual tracking data has consistently failed to show evidence that the controller time-shares among tasks).

The assumption of a fixed number of channels available for tracking and other tasks is another way of saying that the controller's "channel capacity" is constant. Fortunately, we shall not have to determine the total number of channels available, nor shall we have to compute the noise ratio associated with each individual channel. The important point is that these numbers are assumed invariant.

The assumption of white noise processes which scale with signal variance is a direct extension of our model for controller remnant and is consistent with earlier psychophysical data which show that estimation errors tend to scale with the magnitude of the stimulus.

For mathematical convenience we refer central processing sources of interference to the controller's input and treat them as if they were perceptual sources of interference. We thus consider interference to occur among the perceptual tasks required for control of the system. We also assume, however, that position and rate information can be obtained from the same display indicator *without* mutual interference. This assumption is based on manual control data which indicate that the noise ratio associated with estimation of indicator position does not depend on whether or not the controller must also estimate indicator velocity. The noise ratio associated with rate estimation is similarly invariant to the requirement to estimate position (Refs.6 and 8). We therefore associate a single perceptual task with each *indicator* provided by the displays rather than with each *variable* used by the controller. Interference is assumed to occur whether or not the signals presented by the display indicators are linearly correlated.

Equivalent Observation Noise for Single-Variable Tasks

Let us consider a tracking task in which the subject is required to estimate primarily a single variable. We shall derive an expression for equivalent observation noise which shows explicitly the dependence of the noise level upon the number of central-processing channels assigned to the task.

The flow of information is diagrammed in Fig. 1. The displayed variable [x(t)] is corrupted by the noisy, N-channel percaptual pre-processor to yield the perceived variable [y(t)]. Both x(t) and y(t) are presumed to have zero mean. The signal suffers a delay due to limitations of the pilot's neuro-motor system and is then processed by the pilot's equalizer to yield the control signal u(t). Since all sources of randomness have been reflected to the perceptual pre-processor, we need consider here only the relationship between y(t) and x(t).

The output of the n^{th} information-processing channel is

$$y_n(t) = x(t) + r_n(t) \tag{2}$$

where $r_n(t)$ is a Gaussian white noise process having power density level $Q \cdot \sigma_x^2$. The noise ratio Q is assumed to be the same for each channel. Since all N channels are presumed to be allocated to this single task, the total output of the subject's perceptual pre-processor will be

$$y(t) = \sum_{n=1}^{N} [x(t)+r_n(t)] = N\, x(t) + \sum_{n=1}^{N} r_n(t) \tag{3}$$

Noting the linear independence of the $r_n(t)$, we compute the following spectrum for the perceived variable:

$$\begin{aligned} \Phi_{yy} &= N^2 \Phi_{xx} + N\, \sigma_x^2\, Q \\ &= N^2 [\Phi_{xx} + \tfrac{1}{N}\, \sigma_x^2\, Q] \end{aligned} \tag{4}$$

The noise-related portion of the spectrum Φ_{yy} may be reflected to an equivalent observation noise process whose power density level is

$$R^{(1)} = \frac{Q}{N} \sigma_x^2 = P_o \sigma_x^2 \tag{5}$$

where the constant, P_o, replaces Q/N. The unity superscript indicates that this relation holds only when the task is performed alone.

Observation Noise in a Multi-Task Situation

Let us now consider a situation in which M perceptual tasks are performed in parallel. Since the N information channels must now be distributed among the M tasks, the subject will be able to devote only the fraction $f_m^{(M)}$ of his channels to the m^{th} task. The equivalent observation noise is now given as

$$R_m^{(M)} = \frac{Q}{f_m^{(M)} N} \sigma_x^2 = \frac{P_o}{f_m^{(M)}} \sigma_x^2 = P_m^{(M)} \sigma_x^2 \tag{6}$$

We thus show that the primary effect of requiring the subject to perform a multiplicity of tasks is to increase the effective observation noise ratio associated with each component task. The observation noise ratio for the m^{th} task when M tasks are performed simultaneously is simply

$$P_m^{(M)} = P_o / f_m^{(M)} \quad . \tag{7}$$

Transformation of this equation yields

$$f_m^{(M)} = P_o / P_m^{(M)} \quad . \tag{8}$$

Equation (7) may be used for predicting the effects of task interference. Once we determine the observation noise ratio that corresponds to "full capacity", we can predict the increase in noise ratio associated with each display indicator for a particular distribution of channel capacity. The optimal control model then allows us to predict the system performance measures that result from this distribution of capacity.

Equation (8) suggests a measurement technique for determining the pilot's distribution of capacity from the experimental data. For the special case in which the subtasks are independent tracking tasks and can each be characterized by a single variable, we can readily measure the observation noise ratio under single- and multi-axis conditions. If we have confidence in the interference model developed above, we simply equate the fraction of capacity associated with a given subtask to the ratio of the single-axis to the multi-axis observation noise ratio.

Prediction of Multivariable Tracking Performance

Since the model for task interference treats but one aspect of controller behavior, it must be incorporated into a larger modelling system in order that we may predict overall system performance. Indeed, the model represented by Eq. (7) cannot be used to full advantage unless a means exists for predicting the pilot's distribution of capacity among the various subtasks. The optimal control

model discussed in References 7 and 8 provides the required structure for such a modelling system. The following computational procedure has been devised for predicting system performance and pilot behavior in a multivariable control situation in which the pilot's goal is to minimize a quadratic total-task performance measure.

a. Find the set of parameters for the optimal-control model which best reproduces the controller's single-axis behavior. This is a "calibration" procedure which allows one to determine the observation noise ratio that corresponds to full capacity, as well as to determine other basic parameters such as controller time delay, subjective cost functional, and, in the case of peripheral tracking, additional internal noise processes related to peripheral viewing.

b. Compute the total performance cost as a function of the distribution of capacity among the displayed variables, where the fraction of capacity assigned to a given indicator is reflected by an increase in observation noise in accordance with Eq. (8). Since the total capacity must remain fixed, readjustments of the various observation noise ratios must meet the constraint

$$\sum_{m=1}^{M} f_m^{(M)} = \sum_{m=1}^{M} \frac{P_o}{P_m^{(M)}} = 1 \qquad (9)$$

c. Select the capacity distribution that yields the minimum total performance cost. If the total task comprises a set of independent subtasks, pilot behavior and system performance for each subtask may be computed on the basis of the observation noise ratios

associated with the corresponding displays. The amount of "interference" occurring on a given component task may be defined as the performance cost predicted for that task in the M-task situation minus the cost when the task is performed alone.

EXPERIMENTAL VALIDATION

The results of a set of experiments conducted to validate the model are summarized in this section of the paper. The reader is referred to Reference 3 for a complete documentation of the experimentation. Two kinds of control situations were investigated: (a) multiple, independent axes of control, and (b) a single axis of control in which two indicators were displayed to the controller.

Interference Among Independent Axes of Control

The subjects were provided with the four-axis display configuration shown in Fig. 2. Each component display consisted of a moving error bar and a stationary reference line presented on an oscilloscope. This configuration provided four viewing conditions: (a) foveal, (b) 16° peripheral with reference extrapolation possible – as, for example, when fixating the upper left display and tracking a signal on the upper right, (c) 16° peripheral with no reference extrapolation possible, and (d) 22° peripheral, also with no reference extrapolation. Since the stationary reference line became imperceptible a few seconds after peripheral viewing was initiated, the subject's tracking performance was appreciably enhanced whenever he could extrapolate the zero reference from his fixation point to the peripheral display.

A single display was fixated during the entire run length, and two or more axes were controlled simultaneously. Each display used in a given multiaxis experiment was also tracked singly to provide

a set of baseline measures to allow each subject to serve as his own control. Two two-axis manipulators were provided – one controlled by each hand – in order to provide control-display compatibility. K/s plant dynamics were provided on each axis, and each input signal was applied in parallel with the controller's output. Different input waveforms were used on each axis so that the subject would perceive no linear correlations among the inputs. The subjects – all of them instrumented-rated aircraft pilots – were instructed to minimize mean-squared system error when performing a single task and to minimize the sum of the component MS error scores when tracking multiple axes.

We first analyze the results of an experiment in which the subjects were required to fixate the upper left display while tracking all displays simultaneously. No attempt was made to equalize the component task difficulties. On the contrary, since all four input signals were statistically identical, the component task difficulty (in terms of the error score) increased with decreasingly favorable viewing conditions. The performance measures of the four subjects were averaged together for comparison with model predictions.

Model parameters were adjusted to obtain a good match between "predicted" and measured performance measures for each of the four single-axis conditions. See Reference 3 for details of the matching procedure. This matching procedure provides an estimation of the noise ratio corresponding to full capacity. The optimal total multi-task performance score was then computed as follows. First, the optimal-control model was analyzed with several values of observation noise ratio to predict for each axis the relationship between performance score (in this case, tracking error variance) and fraction of capacity. The remaining model parameters were kept

fixed at their nominal values. The predicted relationships are shown for the four viewing conditions in Fig. 3. Optimum four-axis performance was obtained by locating the operating point which yielded minimum total score – defined as the sum of the four component error variance scores – subject to the constraint that the fractions of capacity sum to unity.

Predicted and measured error variance scores are compared in Table 1. The 1-axis and 4-axis scores obtained from the manual control experiments are shown in Table 1a. Also shown are the ratios of the 4-axis to 1-axis scores for each viewing condition and for the total performance measure. Table 1b gives the predicted optimum 4-axis performance: i.e., the performance corresponding to the distribution of capacity that would yield the minimum total score.

Since the subjects were not instructed as to how to apportion the total error among the component scores, the most critical test of the interference model is its ability to predict total score. Table 1 shows that the predicted total performance score was within 3% of the measured score. The model predicts less well the performance on the component axes. The subjects achieved lower scores on the foveal and 22° peripheral tasks than the model would predict, whereas experimental scores were greater than predicted on the axes corresponding to 16° peripheral viewing. These results suggest that the subjects "traded" performance on one pair of component tasks for performance on the remaining two.

The predicted distribution of pilot capacity is shown in the bottom row of Table 1b. The trend of the predictions agrees with our intuitive expectations. The least amount of capacity (10%) was devoted to the easiest (foveal) task, and "attention" increased with component task difficulty up to 45% on the 22° peripheral task.

TABLE 1

Comparison of Measured and Predicted
Error Variance Scores for 4-Axis Experiment

	Measurement	Viewing Condition				Total Score
		Foveal	16° Periph Ref Ext	16° Periph No Ref Ext	22° Periph No Ref Ext	
(a) Measured	1-axis	.11	.25	.42	.96	1.7
	4-axis	.27	.94	1.3	1.6	4.1
	Ratio	2.5	3.8	3.0	1.7	2.4
(b) Predicted: Optimal Behavior	1-axis	.11	.25	.39	.98	1.7
	4-axis	.49	.82	1.1	1.8	4.2
	Ratio	4.6	3.3	2.7	1.9	2.4
	Fract. Cap.	.10	.20	.25	.45	1.00
(c) Predicted: Best Match of Subjects' Behavior	1-axis	.11	.25	.39	.98	1.7
	4-axis	.27	1.0	1.3	1.7	4.2
	Ratio	2.5	4.2	3.2	1.7	2.4
	Fract. Cap.	.20	.15	.20	.50	1.05

Error score in deg^2 visual arc

Average of 4 subjects, 2 trials/subject

Since the subjects' distribution of capacity was apparently different from that predicted by the model, a simple model-matching procedure was used to determine the actual distribution. The curves of Fig. 3 were used to associate a fraction of capacity with the error variance score obtained on each axis when the four axes were controlled together. The results of this procedure are given in Table 1c.* The fractional capacities obtained by this procedure were not subject to the constraint that they sum to unity; nevertheless, we note in Table 1c that the sum of the fractional capacities is, in fact, nearly unity. It appears, then, that the subjects were operating within the constraint of a total fixed capacity.

Model predictions were obtained using the noise ratios shown in Table 1c in order to demonstrate that our model structure accounts for the kinds of 1-axis, 4-axis differences observed in the human controller's frequency-domain measures. (Note that the describing functions and observation noise spectra were not matched by the procedure described above — only the error scores were matched.) Experimental and measured controller describing functions and normalized observation noise spectra** are compared for each of the viewing conditions in Figs. 4-7.

The model predicted the important trends of the differences; namely, the increase in normalized observation noise level, the decrease in controller gain, and the slight increase in high-frequency phase lag as the number of axes tracked was increased from 1 to 4. The most noticeable discrepancy between predicted and measured trends was observed in the observation noise results.

---

*Because we quantized our model results to the nearest integral multiple of 0.05 units of fractional capacity, we could not match the four-axis score perfectly. Comparison of Table 1a and 1c shows, nevertheless, that matching errors were less than 10%.

**Although controller remnant is considered to arise from an underlying *vector* observation noise process, measurements of remnant are presented most readily as an equivalent *scalar* process. Accordingly, controller remnant has been reflected to an equivalent observation noise process added to system error and has been normalized with respect to error variance.

Whereas the experimental data show the 1-axis and 4-axis normalized spectra nearly coinciding at high frequencies, the model indicates that the 1-axis, 4-axis differences should be small at low frequencies and larger at high frequencies. These results may be explained by greater reliance on rate information (relative to position information) than was optimal.* That is, the subjects may have traded position information for rate information within a given axis of control, just as they apparently traded performance on one axis for performance on another, without appreciably affecting their total score. The high-frequency portions of the amplitude-ratio curves shown in Figs. 4-7 are generally higher than those predicted by the model, which is consistent with the notion that the subjects were using somewhat more velocity information than predicted.

Model predictions were also compared with the results of a two-axis tracking experiment. The experimental conditions were basically the same as those described above, except that only two displays were used in a given tracking run: the display which was fixated by the subject and the peripheral display located in the nearest clockwise position. All four pairs of displays were tracked in sequence.

Two-axis performance was predicted on the basis of the single-axis foveal and peripheral measures in the manner described above. Total-task performance was predicted very accurately when the two axes were controlled by different hands (i.e., when the two active displays were separated horizontally). Furthermore, the foveal and peripheral components of the error score were predicted to within ten percent as well. When the two axes were controlled by a single

*A Simple model for controller remnant is presented in Ref. 6 which shows that the break frequency of the normalized observation noise spectrum, as well as the asymptotic low-frequency level, is determined by the ratio of the controller's gain on rate information to his gain on position information.

two-axis manipulator (i.e., when the active displays were separated vertically), the measured scores were about 30% greater than would have been predicted by the model. These results suggest the presence of a significant motor source of interference when a single hand is used for two-dimensional control.

Interference Within a Single Axis

The analyses described above showed that the model for task interference could provide reasonably accurate predictions of the effects of interference in control situations in which the signals displayed to the controller were linearly uncorrelated. An additional experiment was performed to test the validity of the model when the displayed variables were highly correlated. This experiment is documented fully in Reference 3; only the principal results are discussed here.

The subjects were required to control dynamics of the form

$$V = \frac{K}{s(s-1)} \tag{10}$$

These dynamics were implemented as the cascade of two subsystems K_1/s and $K_2/(s-1)$, as shown in Fig. 8. Two display conditions were investigated. For half the training and data trials, only the system error x(t) was displayed. From these data were derived the noise ratio, P_o, as well as other parameters needed to characterize the controller. Both x(t) and the auxiliary signal y(t) – the output of the first subsystem – were displayed for the remaining trials in order to provide a test of the interference model. The display configuration used is shown in Fig. 9. The subjects were instructed to minimize mean-squared system error at all times.

They were not told how to use the auxiliary signal, nor were they assured that the latter signal would aid in their control of system error.

Predictions of two-indicator performance were obtained for the hypotheses of interference and no-interference.* In terms of the model parameters, "no-interference" implied observation noise ratios of P_o associated with all four variables utilized by the controller ($x, \dot{x}, y$, and $\dot{y}$). The hypothesis of interference between the X and Y indicators implied that the noise ratio associated with the joint estimation of x and $\dot{x}$ would be P_o/f_x, and the noise ratio associated with y and $\dot{y}$ would be $P_o/(1-f_x)$, where f_x represents the fraction of pilot capacity allocated to the X indicator.

The hypothesis of interference between the indicators gave the better match to the performance measured when both indicators were displayed to the subjects. The reduction in system error associated with the display of the auxiliary signal $y(t)$ was predicted to within 5%. Changes in the scalar observation noise spectrum, on the other hand, were predicted the least well of all performance measures. As was the case with the multi-axis results, much of the discrepancy between measured and predicted pilot behavior could be accounted for by a slightly nonoptimal distribution of capacity.**

---

*It was not obvious from the data alone whether or not there was interference between the X and Y indicators, since the addition of the Y indicator was expected to reduce mean-squared system error even in the presence of interference. A greater reduction in score was expected in the absence of interference, of course.

**The model predicted that the subject should devote 70% of his attention to the auxiliary signal. Since the subject was required to devote full attention to the X indicator for half of his training runs, it is likely that the mixed training procedure employed induced the subjects to allocate as much capacity as was feasible to the X indicator when both indicators were displayed.

SUMMARY AND DISCUSSION

A model for task interference has been described which is based on the notion that the human controller possesses a fixed amount of capacity which he allocates among the various perceptual tasks to be performed. By relating allocation of capacity to changes in equivalent observation noise ratio, we employed the optimal control model of human behavior to predict total-task performance scores very accurately. In addition, the effects of interference on controller describing functions and observation noise spectra were reasonably well predicted, although less accurately than the effect on total performance score. The ability of the model to account for the trends along all measurement dimensions supported the assumption that interference effects can be related directly to a change in the observation noise ratio.

When multiple axes were tracked simultaneously, the effects of task interference on the component scores were less well predicted than the total performance measure (which was the only measure that the subjects were instructed to regulate). Similarly, the total performance score was predicted very accurately when the subjects performed a single-axis, two-indicator task, but detailed predictions of pilot behavior were less accurate. The major source of this modelling error was attributed to the flexibility permitted the controller in achieving near-optimum system performance. For example, studies performed with the model showed that a somewhat nonoptimal distribution of capacity provided an improved match to a number of performance measures without seriously degrading the accuracy of the predicted total performance score. We would expect to observe this phenomenon when studying realistic flight-control systems, since a well-designed system would ideally be insensitive to moderate variations in the pilot's control and monitoring strategy.

A more serious limitation on the generality of our model for task interference is that we do not yet know how to predict what the pilot's total capacity (in terms of an equivalent observation noise ratio) will be in a given control situation. Although we have found single-axis noise ratios to be about -20 dB for control tasks involving stable vehicle dynamics, a noise ratio of -26 dB was measured when the dynamics were unstable. (The subjects apparently were induced to achieve a particularly low noise level because of the high sensitivity of system performance to observation noise ratio. See Reference 3.) Until we understand better how the pilot's apparent capacity depends upon the nature of the control situation, single-variable "calibration" experiments will be necessary so that nominal model parameter values can be determined. We continue to assume, on the basis of our results, that the pilot's capacity remains fixed for a given type of control situation (and for a given level of pilot proficiency).

Although model and experimental results are in good agreement, the phenomenon of task interference is clearly more complex than our model would indicate. For example, we found that two-axis interference was greater when a single two-axis manipulator is used than when two single-axis manipulators are employed. This result suggests peripheral motor interference effects in addition to the central-processing interference that we have modelled. The near-perfect prediction of the total four-axis score is a bit surprising, then, when one considers that motor interference must have been present in this control situation. Apparently, there were compensating errors in our model structure (e.g., the subjects may have increased their total capacity in this very demanding control situation).

The data against which the model for interference has been tested have been obtained entirely from experiments in which the subjects were not allowed to scan visually. This does not reflect a theoretical limitation on the interference model, however. Consideration of observation noise processes whose statistics vary with time in accordance with the subject's fixation point (see Ref. 7) should enable the optimal-control model to yield predictions of pilot behavior which include the effects both of visual scanning and of central-processing limitations. Further development of the optimal-control model will be necessary before this capability is realized.

Since modern aircraft displays are tending to place multiple display elements in close proximity to one another, our ability to predict the performance of such flight-control systems may depend increasingly less upon our ability to predict scanning behavior and more upon the accuracy with which we can predict mutual interference among elements viewed foveally. We have seen a situation in which the model for task interference correctly predicted the benefits of adding an auxiliary indicator to a display; that is, it predicted that the improvement in performance allowed by the additional information would more than offset the performance degradation caused by interference between display indicators. Accordingly, the author feels that the model for interference may prove useful in the design of complex displays for real flight-control situations, particularly with regard to determining the point at which increased display complexity no longer aids the pilot.

The model for task interference has so far been considered only in terms of its ability to predict multivariable control performance when the subject is presumably working to capacity. While there are good reasons for requiring this type of performance in an experimental

situation, it is highly unlikely that pilots would willingly and routinely fly a vehicle that constantly demands maximum concentration for safe control. Accordingly, the interference model may prove most useful in practice as a technique for predicting how much "attention" is required to meet mission requirements. "Attention" would be associated with observation noise ratio; that is, the greater the noise ratio that could be tolerated, the lower would be the attentional demand of the flight task. In other words, the interference model could be used to predict the relative workload requirements of various tasks. Predictions of workload based on the interference model have been tested against pilot opinion data found in the literature, and good agreement has been found between predicted workload and pilot opinion (Ref.3). Further experimentation is needed to validate the relationship between noise ratio and workload.

CONCLUSIONS

Theoretical and experimental results reviewed in this paper lead to the following conclusions relating to interference among continuous manual control tasks:

(a) The human controller possesses a fixed capacity which must be allocated among the perceptual tasks required.

(b) The direct effect of allocating less than full capacity to a given task is to increase the observation noise ratio associated with that task. The observation noise ratio in this situation is given as P_o/f, where P_o is the observation noise ratio associated with full capacity, and f is the fraction of capacity devoted to the task.

(c) By considering the observation noise ratio as the effective source of controller remnant, one may use an optimal-control model of human behavior to predict changes in pilot describing functions and system performance as a function of fraction of capacity.

(d) The optimal-control model may be used to predict total system performance with great accuracy when independent tracking tasks are performed simultaneously and when multiple displays are used in a single axes of control.

(e) Allocation of capacity and component-task performance can be predicted to the extent that total system performance is sensitive to these parameters. Otherwise, the pilot can be expected to effect some trade-offs among component task scores.

BIBLIOGRAPHY

1. Broadbent, D.E., Perception and Communication, Pergammon Press, New York, 1958.

2. Green, D.M. and Swets, J.A., Signal Detection Theory and Psychophysics, John Wiley and Sons, Inc., New York, 1966.

3. Levison, W.H., Elkind, J.I., and Ward, J.L., "Studies of Multivariable Manual Control Systems: A Model for Task Interference", Rep. No. 1892, (prepared for NASA/Ames Research Center under Contract No. NAS2-3080) Bolt Beranek and Newman Inc., Cambridge, Mass., December 8, 1969.

4. McRuer, D.T., Graham, D., Krendel, E.S., and Reisener, W., Jr., Human Pilot Dynamics in Compensatory Systems Theory, Models and Experiments with Controlled-Element and Forcing Function Variations, Air Force Flight Dynamics Laboratory, Wright-Patterson Air Force Base, Ohio, AFFDL-TR-65-15, July 1965.

5. Levison, W.H., Kleinman, D.L., and Baron, S., A Model for Human Controller Remnant (Final Report), Rep. No. 1731, (prepared for NASA/G.C. Marshall Space Flight Center, Huntsville, Alabama under Contract No. NAS8-21136) Bolt Beranek and Newman Inc., Cambridge, Mass., October 15, 1968.

6. Levison, W.H., Baron, S., and Kleinman, D.L., "A Model for Human Controller Remnant," IEEE Trans. on Man-Machine Systems, MMS-10:101-108, December 1969.

7. Baron, S. and Kleinman, D.L., "The Human as an Optimal Controller and Information Processor," IEEE, Trans. on Man-Machine Systems, MMS-10:9-17, March 1969.

8. Kleinman, D.L., Baron, S., and Levison, W.H., "An Optimal Control Model of Human Behavior," Proceedings of the Fifth Annual NASA/University Conference on Manual Control, M.I.T., Cambridge, Mass., March 27-29, 1969.

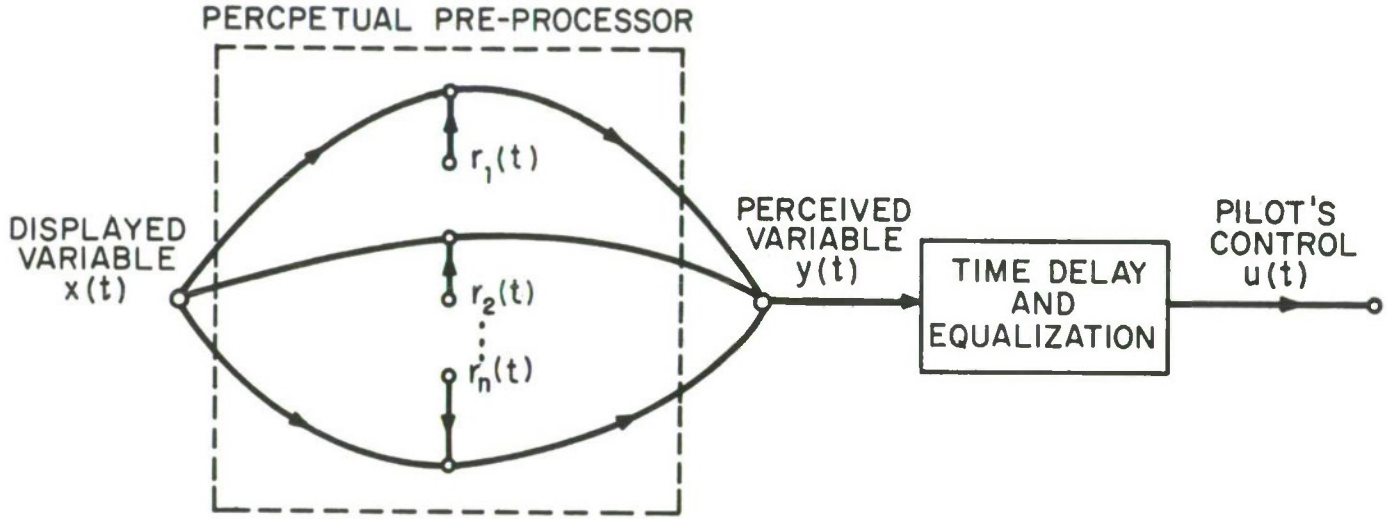


FIG.1 FLOW OF INFORMATION THROUGH THE HUMAN CONTROLLER FOR A SINGLE-VARIABLE TRACKING SITUATION

WHL-034

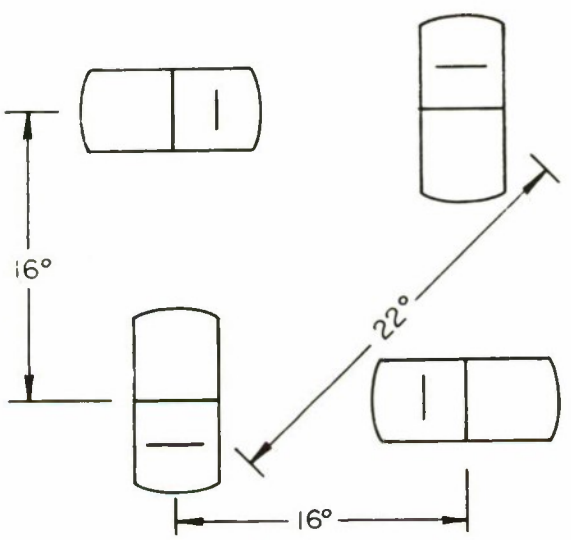


FIG.2 DISPLAY CONFIGURATION USED IN THE EXPERIMENTS

Dimensions Shown in Degrees of Visual Arc

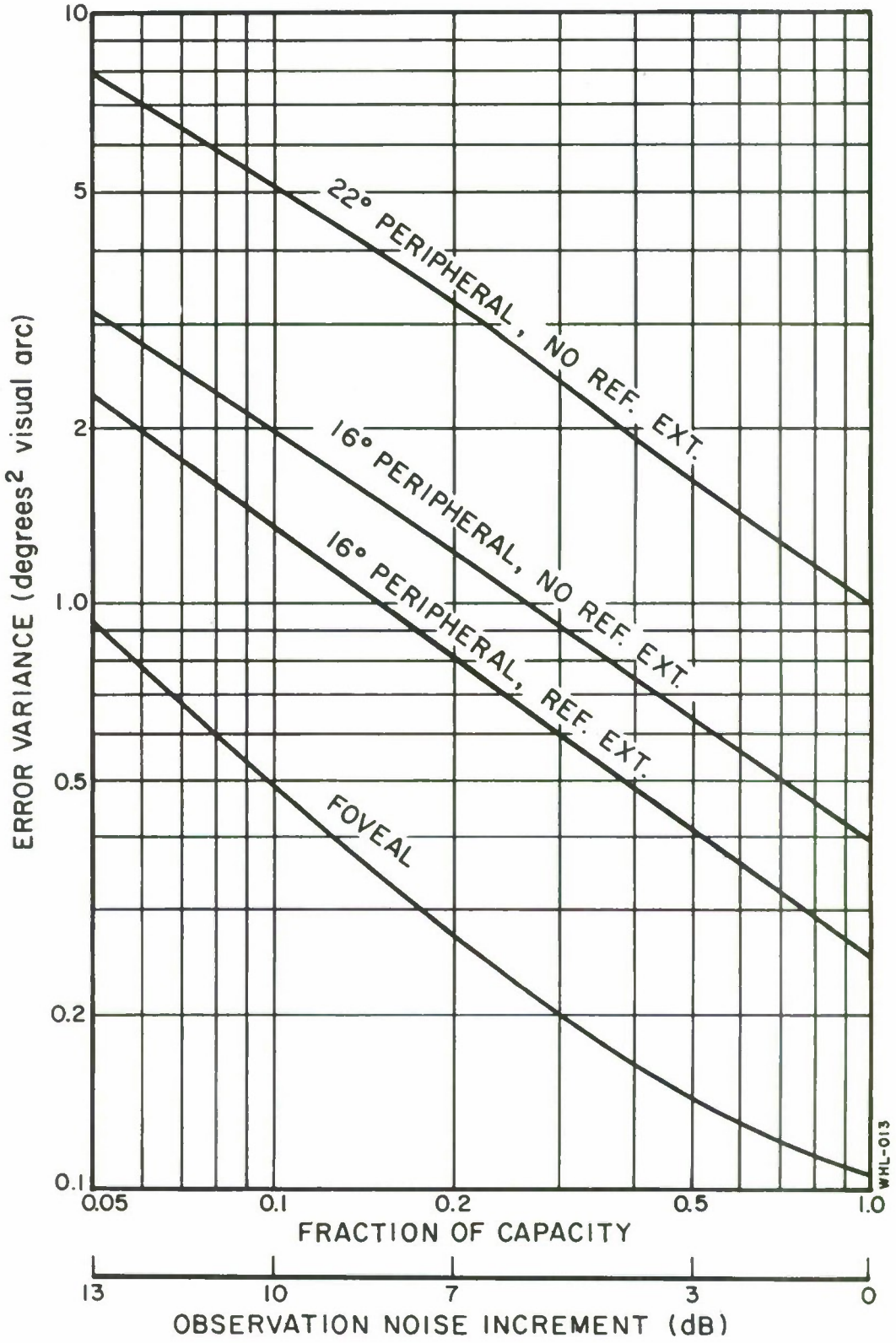


FIG.3 PREDICTED RELATIONSHIP BETWEEN ERROR VARIANCE AND FRACTIONAL ALLOCATION OF CAPACITY

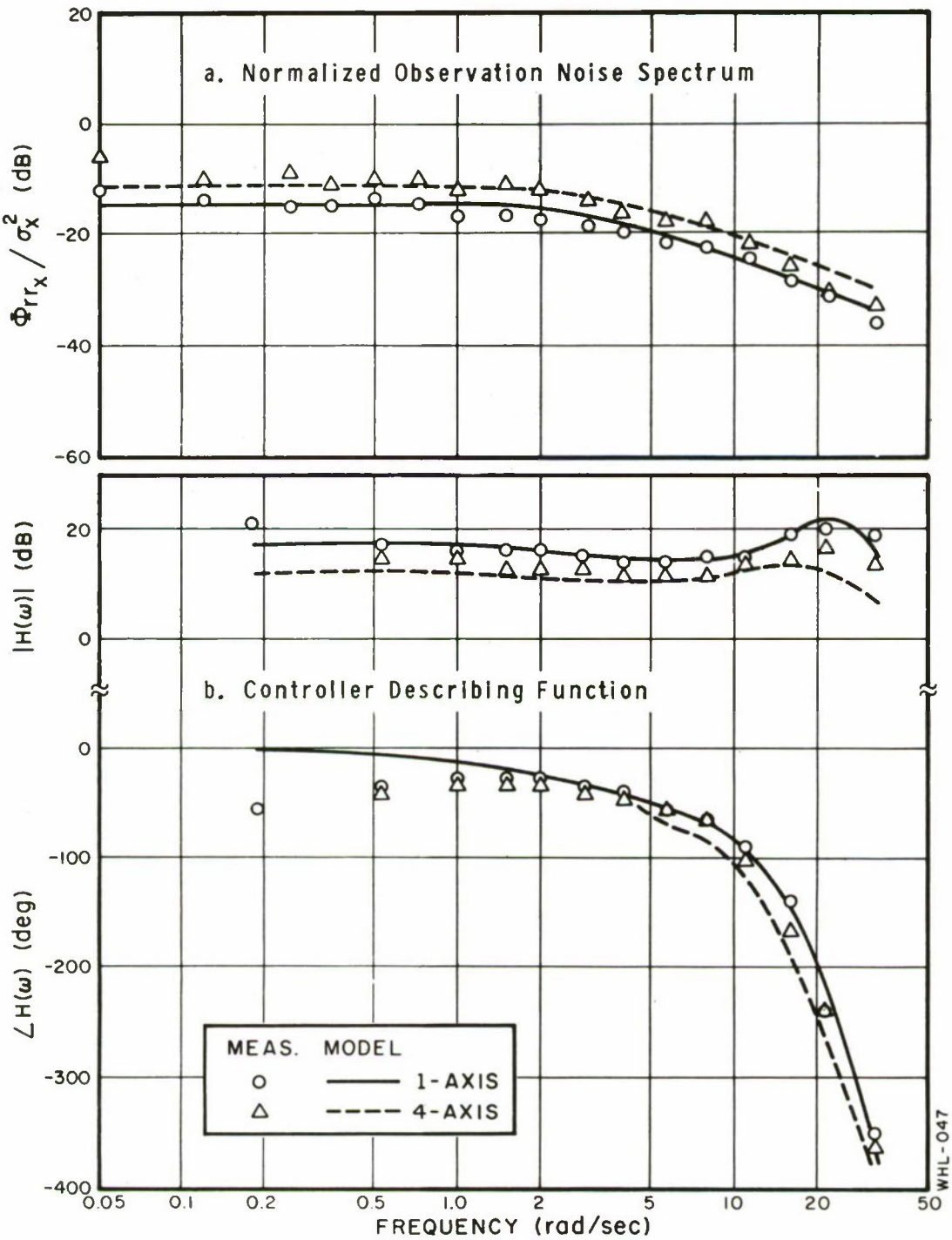


FIG.4 EFFECT OF NUMBER OF AXES TRACKED ON FREQUENCY-DOMAIN MEASURES: FOVEAL VIEWING

Average of 4 subjects, 2 trials/subject

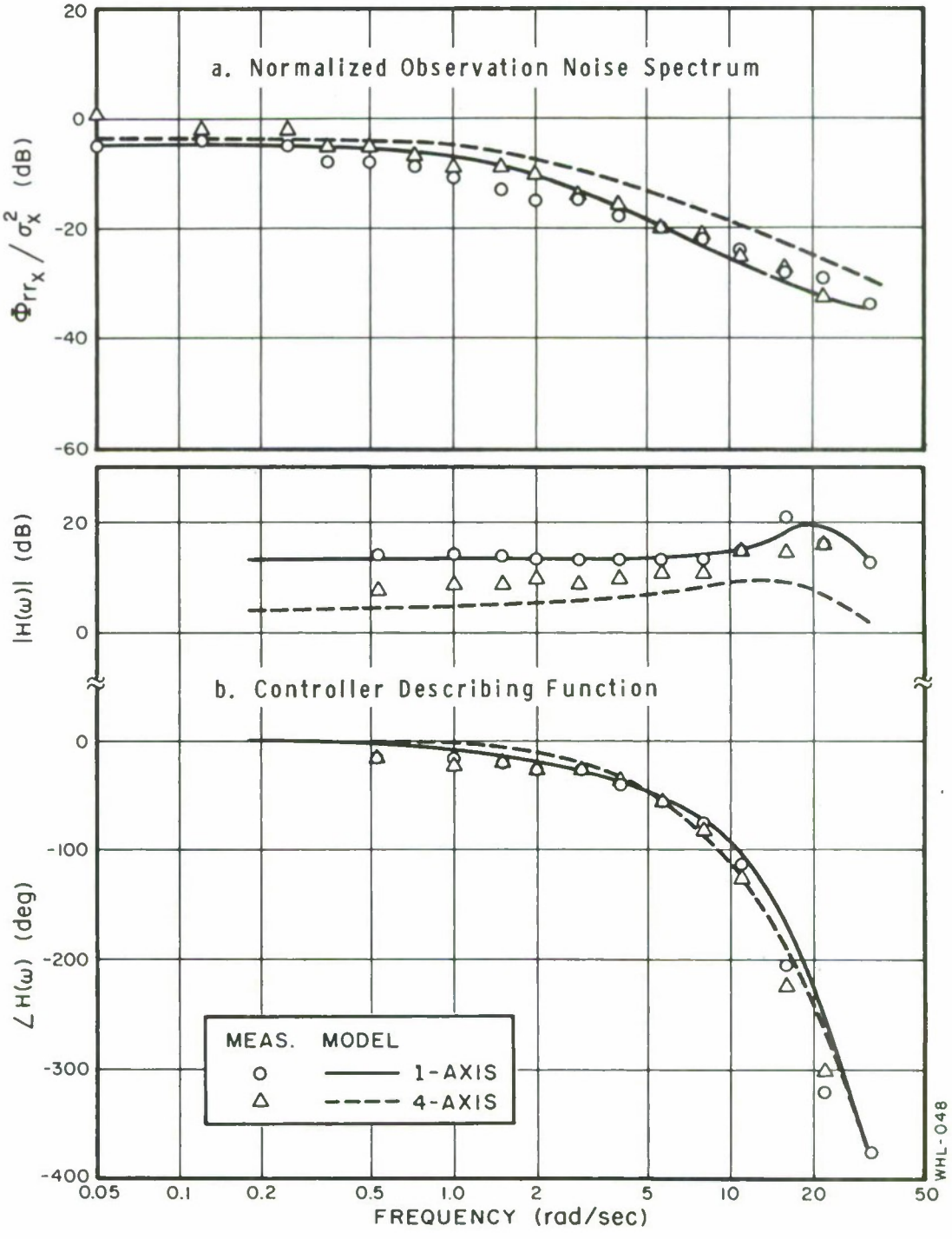


FIG.5 EFFECT OF NUMBER OF AXES TRACKED ON FREQUENCY-DOMAIN MEASURES: 16° PERIPHERAL VIEWING WITH REFERENCE EXTRAPOLATION

Average of 4 subjects, 2 trials/subject

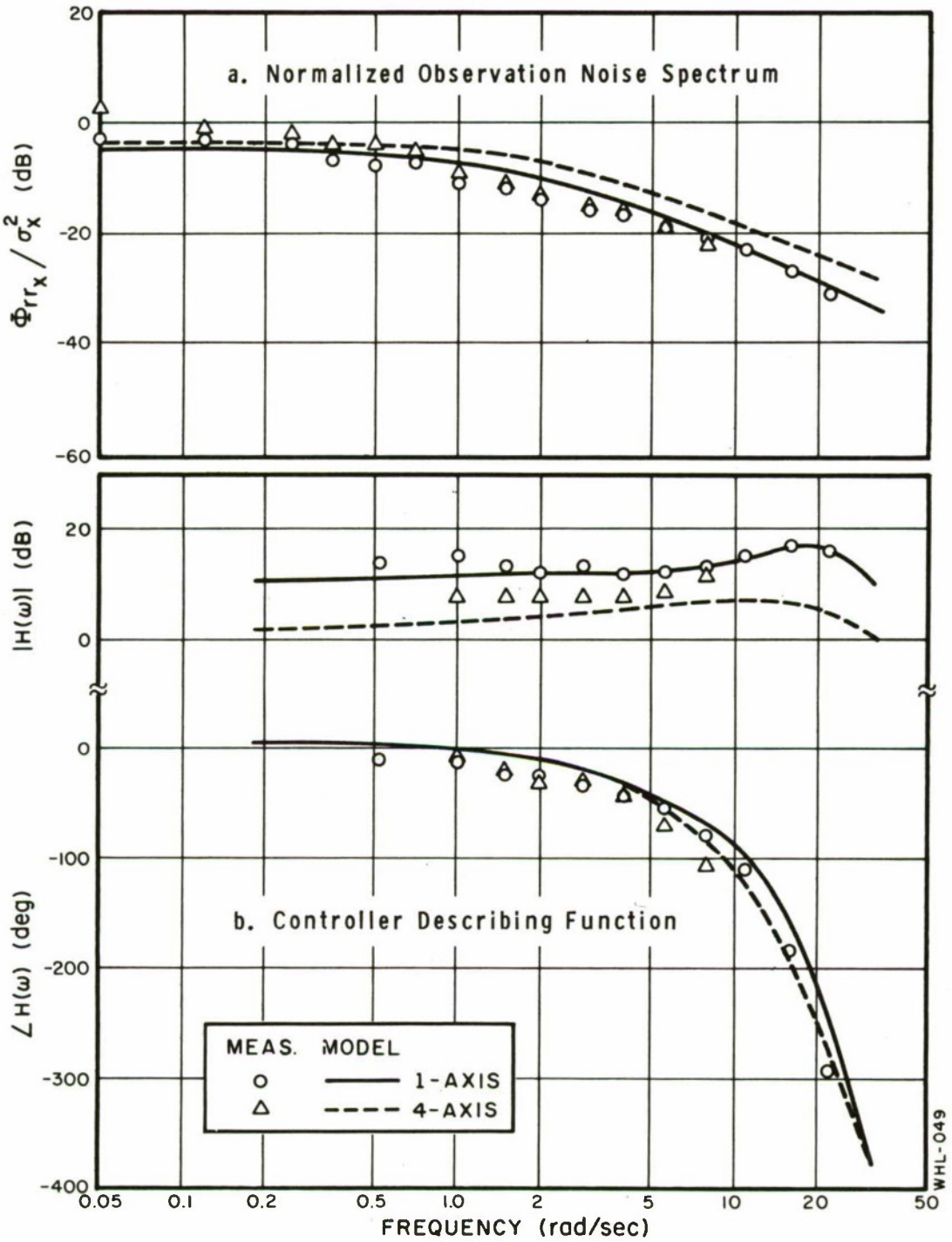


FIG.6 EFFECT OF NUMBER OF AXES TRACKED ON FREQUENCY-DOMAIN MEASURES: 16° PERIPHERAL VIEWING WITHOUT REFERENCE EXTRAPOLATION

Average of 4 subjects, 2 trials/subject

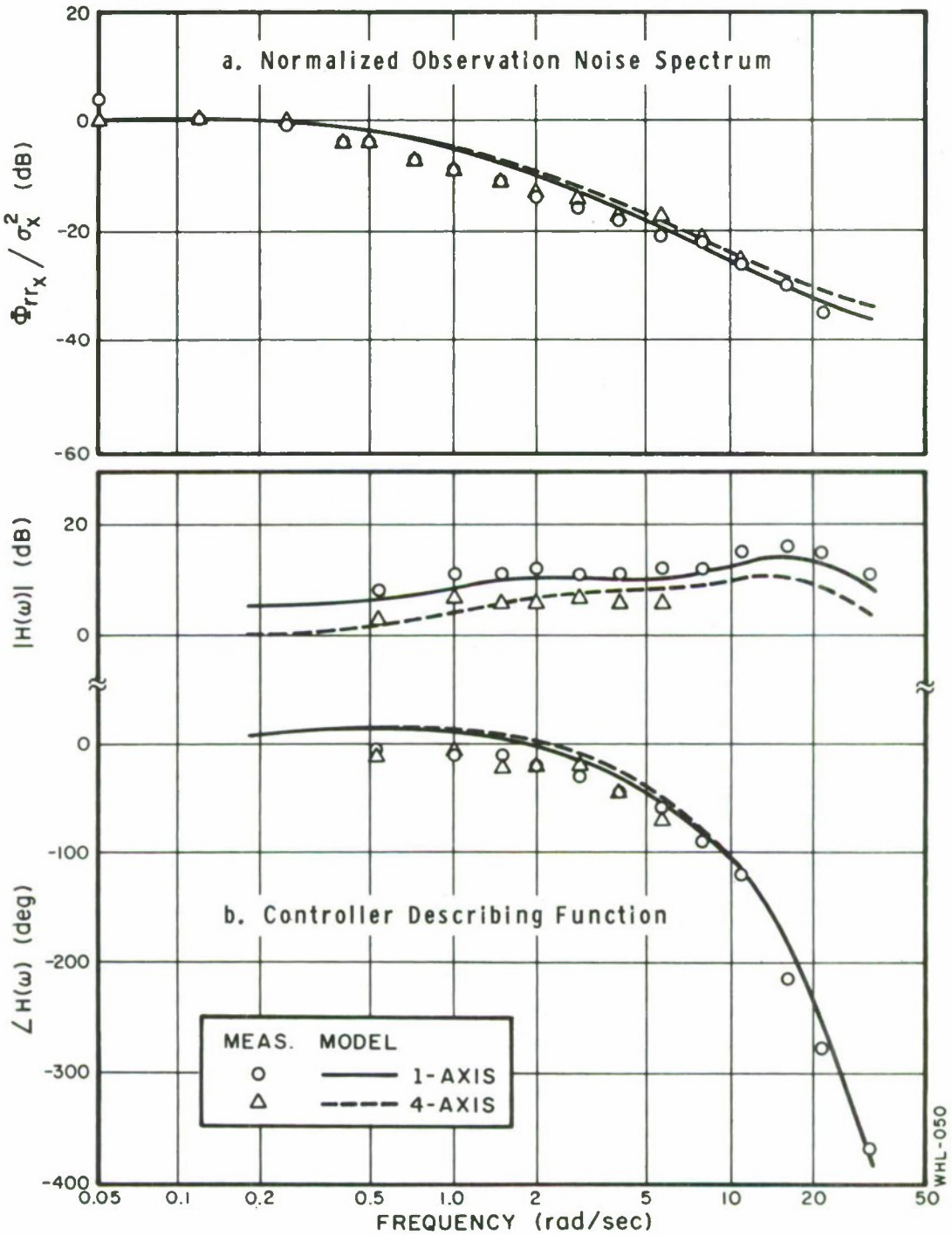


FIG.7 EFFECT OF NUMBER OF AXES TRACKED ON FREQUENCY-DOMAIN MEASURES: 22° PERIPHERAL VIEWING WITHOUT REFERENCE EXTRAPOLATION

Average of 4 subjects, 2 trials/subject

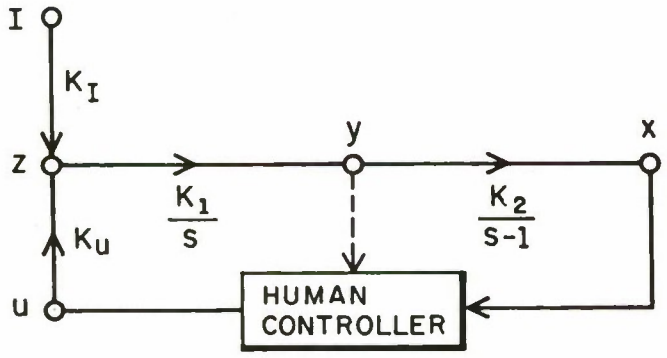


FIG.8 DIAGRAM OF THE SINGLE-AXIS, MULTIVARIABLE CONTROL SITUATION

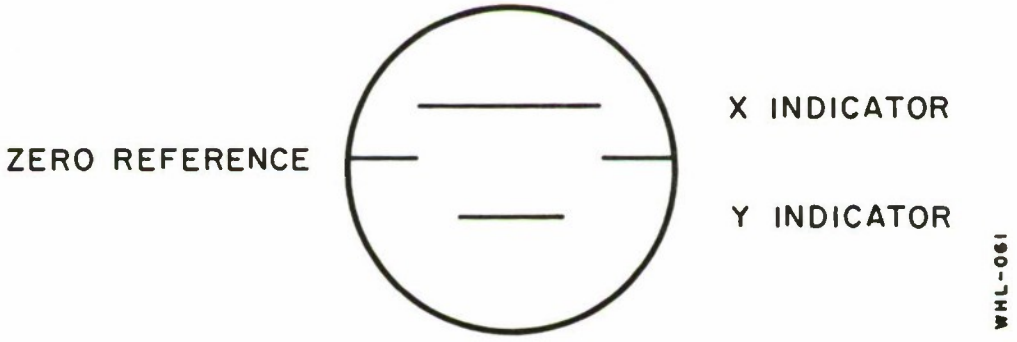


FIG. 9 EXPERIMENTAL DISPLAY FORMAT

On How Often the Supervisor Should Sample*

Thomas B. Sheridan**

18 December 1969

ABSTRACT

This paper presents a procedure for specifying how long a supervisor or monitor of a process should wait between input samples to maximize a given value or payoff function, assuming he resets the controls with each sample as a function of the best information he has. The procedure is based upon Bayesian preposterior information analysis.

*The research reported in this paper was supported by NASA Grant NGR 22-009-002. The author acknowledges the assistance of William B. Rouse and William H. Vickers in implementing calculations for the examples.

**Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

INTRODUCTION

The supervisor of a process is here defined as a person or function which adjusts a control variable y in an effort to maximize a given value function V of the system's independent input x, control input y, and output z. A typical diagram is given in Fig. 1. Often this is called a "learning" or "optimizing" control system. It is easiest to think of V as a continuous earnings rate (equivalent to money per unit time) and of x, y and z as continuous in time.

Sometimes, as with a steepest ascent optimizer or an operant conditioned learner, z information is not available to the supervisor and the dependence of V on x and y must be gained from experience. Sometimes z is available to the supervisor, and either by being told or after sufficient experience the functions $z(x,y)$ and $V(x,y,z)$ become known independently. However, if $z(x,y)$ is deterministic and known to the supervisor, it is also redundant to $V(x,y,z)$. z can then be treated as an "intervening variable" either of no direct interest to the supervisor, or in any case not necessary in determining $V(x,y)$, and the system reduces to the simpler form of Fig. 2.

Fig. 3 caricatures the supervisor sampling the input variable and adjusting the control variable to maximize the flow of value (happiness, money) from the process. Fig. 4 defines the problem: what is given and what is sought.

At one extreme the supervisor may monitor x and V continuously. In this case we assume, as the perfect supervisor, he will continuously modify y to maximize V. At the other extreme he may refrain from monitoring x and may set y to an optimum on the basis of some distributional constraints on x and V gained from past experience. Finally, he may adapt a mixed strategy. The contribution of this paper is to call attention to the nature of this mixed strategy. The question which concerns us here is how often should x be sampled and how is this interval related to his control function between samples as well as the cost of sampling.

THEORY

Before considering the sampling rate optimization itself let us consider the two extremes: 1) continuously observing each new value of x and adjusting y to maximize value (earnings rate) $V(x,y)$ when x is known; 2) never observing x and adjusting y to maximize V on the basis of a prior probability distribution of x. The theory used to analyze these two

extremal situations is derived from so-called "preposterior" analysis or "information value theory," as presented by Howard [1] and others.

We start with the following assumptions and notation:

x is a random event with known probability density $\{x\}$.

$V(x,y)$, the value gained per unit time, has known expected value $\langle V \mid xy \rangle$ when x and y are specified.

Then the expected value of $V(y)$ is

$\langle V \mid y \rangle = \int_x \langle V \mid xy \rangle \{x\}$ where $\int$ is a generalized summation over all x. (1)

1. Knowledge of Input Only by Prior Distribution

If the supervisor knows only the a priori distribution $\{x\}$ then his best strategy is to determine $\langle V \mid y \rangle$ and adjust y to maximize this function, yielding the overall expected value of V in this case,

$$\langle V_1 \rangle = \max_y \langle V \mid y \rangle = \max_y \int_x \langle V \mid xy \rangle \{x\}. \quad (2)$$

2. Perfect Knowledge of Input

If the supervisor has perfect knowledge of each x then his best strategy is to adjust y to maximize $\langle V \mid xy \rangle$ for each successive x.

$$\langle V_2 \mid x \rangle = \max_y \langle V \mid xy \rangle. \quad (3)$$

Before actually knowing each x he can say that <u>if</u> he did know x <u>and</u> he were to optimize y with respect to each x, <u>and</u> since x occurs according to the known prior distribution $\{x\}$, then he can be sure that the expected value of V afterwards would be

$$\langle V_2 \rangle = \int_x \langle V_2 \mid x \rangle \{x\} = \int_x \left[\max_y \langle V \mid xy \rangle\right] \{x\}. \quad (4)$$

It is important to note that this can be stated without knowing what x actually is at any particular time, just that it is known and that the optimal control strategy is followed. This is what was previously referred to as preposterior analysis.

3. Value of Continuous Information

$\langle V_2 \rangle$ will always be greater than $\langle V_1 \rangle$ and their difference may be considered the expected value of the continuous supervision. Howard [1]

has called this the "value of clairvoyance."

4. Intermittent Sampling of Input with Continuous Control Variable Adjustment Between Samples

The supervisor need not operate at either of the extremes, continuous observation of x and adjustment of y or once-and-for-all setting of y. It is more often the case that knowledge of x can be updated periodically or whenever some threshold of uncertainty is reached. This assumes x at time t after sampling correlates with x sampled.

Now consider a variable x' representing the supervisor's state of knowledge about x at some intermediate time between the two extremes cited above, at the one extreme knowing x exactly and at the other extreme knowing x only as a stationary distribution $\{x\}$. Thus x' will be a nonstationary random variable with probability density function $\{x' \mid x_o t\}$, a function of time after a sampled and measured value $x = x_o$ at $t = 0$.

$$\{x' \mid x_o t\}_{t=0} = \text{an impulse at } x_o \tag{5}$$

$$\{x' \mid x_o t\}_{t=\infty} = \{x\} \tag{6}$$

In between, $\{x' \mid x_o t\}$ can be assumed to have monotonically increasing variance and, assuming $\{x\}$ has a mean of zero, $\{x' \mid x_o t\}$ has a mean which systematically moves from x_o to 0 as t goes from 0 to ∞. This is illustrated in Fig. 5.

We assume availability of some knowledge or model of the secular changes of $\{x' \mid x_o t\}$. Alternative models are discussed below, in Section 6. First consider the optimum strategy of a supervisor who knows $\{x' \mid x_o t\}$.

The supervisor's best strategy at each t (in discrete or continuous time) is to optimize y based upon $\{x' \mid x_o t\}$ analogous to what he would do in equation (2) where he knows only $\{x\}$; but $\{x' \mid x_o t\}$ gives him better information about x. His best estimate of value V_3 at this intermediate t, for the given x_o and for alternative values of y he might choose to apply, is obtained by averaging over the known $\{x' \mid x_o t\}$ as in (1):

$$\langle V_3 \mid x_o y t \rangle = \int_{x'} \langle V \mid x' y \rangle \{x' \mid x_o t\} \tag{7}$$

where $\langle V \mid x' y \rangle = \langle V \mid x y \rangle$. After selecting y to maximize $\langle V_3 \rangle$ for the

given x_o and t the supervisor has

$$\langle V_3 \mid x_o t \rangle = \max_y \langle V_3 \mid x_o y t \rangle = \max_y \int_{x'} \langle V \mid x' y \rangle \{x' \mid x_o t\}. \tag{8}$$

In a manner analogous to (4), the supervisor knows that if he knew x_o and t and were to select y optimally at each t based upon an assumed model for $\{x'\}$, he can specify his expected return before the fact on the basis of the distribution of x_o,

$$\langle V_3 \mid t \rangle = \int_{x_o} \langle V_3 \mid x_o t \rangle \{x_o\}. \tag{9}$$

But x_o has the same prior distribution as x,

$$\{x_o\} = \{x\}, \tag{10}$$

which is known. Therefore the supervisor can determine before the fact of knowing any particular x_o that his expected earn rate at time t after a sample is

$$\langle V_3 \mid t \rangle = \int_x \langle V_3 \mid xt \rangle \{x\} = \int_{x_o} \left[\max_y \int_{x'} \langle V \mid x' y \rangle \{x' \mid x_o t\} \right] \{x_o\} \tag{11}$$

This is the expected instantaneous rate of return to be obtained from continuous control t seconds after the sample. Note that, because of (5), when $t = 0$ (11) reduces to (4). When $t = \infty$, the random variable x' becomes independent of x_o; in this case derivation of (11) need go no further than (8), or equivalently the expectation over $\{x_o\} = \{x\}$ is deleted from equation (11) reducing it to the form of (2).

As t increases $\langle V_3 \mid t \rangle$ will become less like $\langle V_2 \rangle$ and more like $\langle V_1 \rangle$. With increasing t one is less able to select a y suited for a narrow range of x, for x' becomes a worse estimate of the unknown x. As $t \to \infty$ $\{x' \mid x_o t\} \to \{x\}$.

5. Selection of Optimum Sampling Interval

The expected value of return per-unit-time averaged over a sampling interval T is

$$\langle V_3 \rangle_T = \frac{1}{T} \int_{t=0}^{T} \langle V_3 \mid t \rangle. \tag{12}$$

In order to decide how often to sample, the inherent cost of sampling must

also be considered. If, for example, the cost-per-sample is C, then over a sampling interval T the cost of that sample per unit time would be C/T. The expected net value per-unit-time of sampling after each interval T would then be

$$\left\langle V_{net} \right\rangle_T = \left\langle V_3 \right\rangle_T - \frac{C}{T} \tag{13}$$

The optimum strategy therefore is to pick the sampling interval T which maximizes $\left\langle V_{net} \right\rangle_T$. This is indicated in Fig. 6.

6. Intermittent Sampling of Input with Control Variable Adjustment Only at the Time of Sample

In sections 5 and 6 above it was assumed that the supervisor incurred a significant cost when taking samples of input x, but that no significant cost was incurred when control variable y was changed. If we assume that only when a sample is taken can the control variable be adjusted, and that a combined cost C' is incurred whenever this happens, equations 8 through 13 above must be modified. Earnings per-unit-time in this system will be called V'.

Equation (7) obtains as before. But in the present case y cannot be maximized separately at each t, but only at t=0 when the sample x_o is taken. Thus first we must average value over the interval,

$$\left\langle V_3' \mid x_o y \right\rangle_T = \frac{1}{T} \int_{t=0}^{T} \left\langle V_3 \mid x_o y t \right\rangle \tag{14}$$

With respect to this time-averaged expectation over the interval we optimize y

$$\left\langle V_3' \mid x_o \right\rangle_T = \max_y \left\langle V_3' \mid x_o y \right\rangle_T \tag{15}$$

and then expect over the probable values of x_o, where $\{x_o\} = \{x\}$

$$\left\langle V_3' \right\rangle_T = \int_{x_o} \left[\max_y \left\langle V_3' \mid x_o y \right\rangle_T \right] \{x_o\} \tag{16}$$

The expected net value per-unit-time of both sampling x and adjusting y at intervals of T is then

$$\left\langle V_{net}' \right\rangle_T = \left\langle V_3' \right\rangle_T - \frac{C'}{T} \tag{17}$$

7. Expected Behavior of Input Following a Sample

The form of $\{x' \mid x_o t\}$ depends upon: 1) the bandwidth of $x(t)$; 2) the order of the sample measurement, i.e., the number of derivatives measured. From the sampling theorem we know that one zero order (position) sample every $1/2\omega_c$, where ω_c is frequency bandwidth in hertz, is sufficient to reconstitute the signal perfectly, provided a large number of samples have been taken. There is a sizeable body of theory pertinent to predicting stationary time series, much of it due to Weiner [2].

Predictive models have been applied previously to problems of human behavior, but to the writer's knowledge, they have not been applied to the specific problem of preposterior optimization of sampling rate assuming optimization of the control setting based upon such sampling. Smallwood [3] and Carbonell [4] have suggested predictive models for a human operator's knowledge about the reading of an instrument as a function of time since sampling. Smallwood called such an $\{x' \mid x_o t\}$ an "internal model." His models were linear time invariant filters operating upon white Gaussian noise. Carbonell's model assumed multiple input functions (like aircraft instrument readings) which queue up in competition for attention samples meted out by the supervisor. Carbonell's queueing theoretical approach depends on the last readings of instruments, time elapsed since each last reading, and a risk-cost heirarchy. Sheridan [5] and others have applied simple predictive internal models to human control behavior with sampling considerations.

It is of utmost importance that the assumed model for $\{x' \mid x_o t\}$ have a reasonably correct relation between the mean and variance or range as a function of time. The model must not be chosen arbitrarily; it is possible for an assumed $\{x' \mid x_o t\}$ which initially appears reasonable to yield a $\langle V_3 \mid t \rangle$ which is unreasonable. For example it can be shown that some assumed distributions will lead to $\langle V_3 \mid t \rangle$ functions which are not monotonically decreasing with time since the last sample.

EXAMPLE FOR AN ARBITRARY VALUE FUNCTION

Let us assume a supervisor is attending a process forced by a random variable x, whose probability density function is Gaussian with mean 0 and variance v_o. Let the value function (and/or the expected value function)

$V(x,y) = \langle V \mid xy \rangle$ be given by $(K_1x^2y - K_2y^2)$. The latter function, while arbitrarily chosen here to ensure a finite optimum value of y, might occur in asymetric control situations where there is utility in a positive control correlation with large absolute values of input while cost of control is assumed proportional to y^2 (analogous to power used).

Let us assume for simplicity that the process generating x is equivalent to the output of a first order filter $1/(1+s)$ operating on white Gaussian noise. This is a reasonable time series model for a variety of real world problems and was actually used by Smallwood [3] to characterize aircraft instrument signals. If the distribution of x, which for this process is necessarily Gaussian, has mean of zero and variance v_o then, as pointed out by Smallwood, our state of knowledge of x at some time t after a sample x_o is represented by a distribution of x' having mean

$$m'(t) = x_o \exp(-At) \tag{18}$$

and variance

$$v'(t) = v_o[1 - \exp(-2At)] \tag{19}$$

so that

$$\{x' \mid x_o t\} = \frac{1}{\sqrt{2\pi v'}} \exp\left(-\frac{(x'-m')^2}{2v'}\right) \tag{20}$$

The numbers in the development below correspond to those in the previous section:

1. In the worst case where the supervisor knows only the prior distribution x, his best estimate of $\langle V_1 \mid y \rangle$ is

$$\langle V_1 \mid y \rangle = \int_x \langle V \mid xy \rangle \{x\} = \int_{-\infty}^{\infty} (K_1x^2y - K_2y^2) \frac{1}{\sqrt{2\pi v_o}} \exp\left[-\frac{x^2}{2v_o}\right] dx$$

$$= K_1yv_o + K_2y^2 \tag{21}$$

To find $\langle V_1 \rangle$ he adjusts y once and for all to maximize $\langle V_1 \mid y \rangle$:

$$\frac{\partial \langle V_1 \mid y \rangle}{\partial y} = \frac{\partial (K_1yv_o - K_2y^2)}{\partial y} = K_1v_o - 2K_2y = 0 \tag{22}$$

so he sets $y = \frac{K_1 v_o}{2K_2}$. Then by substitution for y his overall expected value

$$\langle V_1 \rangle = \max_y \langle V_1 \mid y \rangle = \frac{K_1^2 v_o^2}{2K_2} - \frac{K_1^2 v_o^2}{4K_2} = \frac{K_1^2 v_o^2}{4K_2} \tag{23}$$

2. If the supervisor has perfect knowledge of each x then he adjusts y to maximize $\langle V \mid xy \rangle$ continuously for each successive x.

$$\frac{\partial \langle V \mid xy \rangle}{y} = \frac{\partial (K_1 x^2 y - K_2 y^2)}{\partial y} = K_1 x^2 - 2K_2 y = 0 \tag{24}$$

so he sets $y = \frac{K_1 x^2}{2K_2}$. Notice that his adjustment of y is a function of x, for we cannot know what x value is observed a priori. Then

$$\langle V_2 \mid x \rangle = \max_y \langle V \mid xy \rangle = K_1 x^2 \left[\frac{K_1 x^2}{2K_2}\right] - K_2 \left[\frac{K_1 x^2}{2K_2}\right]^2$$

$$= \frac{K_1^2}{2K_2} x^4 - \frac{K_1^2}{4K_2} = \frac{K_1^2 x^4}{4K_2} . \tag{25}$$

Now the overall expected value $\langle V_2 \rangle$ is found by taking the expectation of $\langle V_2 \mid x \rangle$ over x,

$$\langle V_2 \rangle = \int_x \langle V_2 \mid x \rangle \{x\} = \int_{-\infty}^{\infty} \frac{K_1^2 x^4}{4K_2} \frac{1}{\sqrt{2\pi v_o}} \exp[-\frac{x^2}{2v_o}] \, dx$$

$$= \frac{3K_1^2 v_o^2}{4K_2} . \tag{26}$$

3. It is seen that

$$\langle V_2 \rangle - \langle V_1 \rangle = \frac{3K_1^2 v_o^2}{4K_2} - \frac{K_1^2 v_o^2}{4K_2} = \frac{K_1^2 v_o^2}{2K_2} \tag{27}$$

This is the additional value of knowing each x as it comes, over and above knowing only $\{x\}$. Note that in this example the value of continuously

updated information is twice the value of prior information by itself.

4. If the supervisor samples x intermittently and can put in effect a policy of continuously adjusting y based upon a model of what x will do following the taking of a sample x_o and until the next sample, as described earlier in section 4 of the theoretical exposition, the value is specified within the interval $0 < t < T$ by equation (11). Working up to (11) in steps, (7) is first applied. For our assumed model for $\{x' \mid x_o t\}$, equations (18) through (20),

$$\left\langle V_3 \mid x_o y t \right\rangle = \int_{-\infty}^{\infty} (K_1 x'^2 y - K_2 y^2) \frac{1}{\sqrt{2\pi v'}} \exp\left[-\frac{(x' - m')^2}{2v'}\right] dx'$$

$$= K_1 y(v' + m'^2) - K_2 y^2. \tag{28}$$

The above integral is most easily solved by substituting a dummy variable for $(x' - m')$.

In terms of this information an optimum y is determined for x_o at each t

$$\frac{\partial}{\partial y} \left\langle V_3 \mid x_o y t \right\rangle = \frac{\partial}{\partial y} [K_1 y(v' + m'^2) - K_2 y^2]$$

$$= K_1(v' + m'^2) - 2K_2 y = 0, \tag{29}$$

y is then set to $\dfrac{K_1(v' + m'^2)}{2K_2}$.

After substitution of this expression for y we obtain

$$\left\langle V_3 \mid x_o t \right\rangle = \max_y \left\langle V_3 \mid x_o y t \right\rangle = \frac{K_1^2(v' + m'^2)^2}{2K_2} - \frac{K_1^2(v' + m'^2)^2}{4K_2}$$

$$= \frac{K_1^2(v' + m'^2)^2}{4K_2} \tag{30}$$

where m' and v' are functions of x_o, v_o and t. Considering the distribution of possible values for x_o and evaluating m' and v' in terms of (18) and (19),

$$\langle V_3 \mid t \rangle = \int_{-\infty}^{\infty} \langle V_3 \mid x_o t \rangle \{x_o\}\, dx_o$$

$$= \int_{-\infty}^{\infty} \frac{K_1^2}{4K_2} \left[v_o - v_o \exp(-2At) + x_o^2 \exp(-2At) \right]^2 \frac{1}{\sqrt{2\pi v_o}} \exp\left[-\frac{x_o^2}{2v_o} \right] dx_o$$

$$= \frac{K_1^2 v_o^2}{4K_2} \left[1 + 2 \exp(-4At) \right] \tag{31}$$

Checking the limiting conditions we find, as expected,

$$\langle V_3 \mid t = 0 \rangle = \frac{3K_1^2 v_o^2}{4K_2} = \langle V_2 \rangle . \tag{32}$$

$$\langle V_3 \mid t = \infty \rangle = \frac{K_1^2 v_o^2}{4K_2} = \langle V_1 \rangle . \tag{33}$$

It is easy to assume a model for $\{x' \mid x_o t\}$ with what seem to be reasonable properties which will yield rather unreasonable values of $\langle V_3 \mid t \rangle$. Any reasonable model must conform to the multiplicative probability constraint that

$$\sum_{x_1} \left\{ p \left[\begin{array}{l} x=x_2 \\ \text{at } t=t_2 \end{array} \middle| \begin{array}{l} x=x_1 \\ \text{at } t=t_1 \\ \text{and} \\ x=x_o \\ \text{at } t=t_o \end{array} \right] \cdot p \left[\begin{array}{l} x=x_1 \\ \text{at } t=t_1 \end{array} \middle| \begin{array}{l} x=x_o \\ \text{at } t=t_o \end{array} \right] \right\} = p \left[\begin{array}{l} x=x_2 \\ \text{at } t=t_2 \end{array} \middle| \begin{array}{l} x=x_o \\ \text{at } t=t_o \end{array} \right] \tag{34}$$

which prevents an arbitrary choice for functions of central tendency and variability for x' and constrains their dependence upon t.

5. To determine $\langle V_3 \rangle_T$, the expected return averaged over a sampling interval T, we calculate

$$\langle V_3 \rangle_T = \frac{1}{T} \int_{t=o}^{T} \langle V_3 \mid t \rangle \, dt = \frac{K_1^2 v_o^2}{4K_2} \int_o^T \left[1 + 2 \exp(-4At) \right] dt$$

$$= \frac{K_1^2 v_o^2}{4K_2} \left[1 + \frac{1}{2AT} - \frac{\exp(-4AT)}{2AT} \right] . \tag{35}$$

After subtracting the average sampling cost over the interval $\frac{C}{T}$ the net averaged expected return over the sampling interval is

$$\left\langle V_{net} \right\rangle_T = \frac{K_1^2 v_o^2}{4K_2} \left[1 + \frac{1}{2AT} - \frac{\exp(-4AT)}{2AT} \right] - \frac{C}{T} . \tag{36}$$

This is the kind of situation illustrated in Fig. 4.

6. If, in accordance with section 6 of the theoretical exposition, the supervisor can only adjust y when he samples x, and he incurs a combined cost C' for each combined sample-plus-adjust action, then the above example must follow the development of equations 14-17.

$$\begin{aligned} \left\langle V_3' \mid x_o y \right\rangle_T &= \frac{1}{T} \int_{t=o}^{T} \left\langle V_3 \mid x_o y t \right\rangle = \frac{1}{T} \int_{t=o}^{T} \left[K_1 y (v' + m'^2) - K_2 y^2 \right] \\ &= \frac{1}{T} \int_{t=o}^{T} \left[K_1 y [v_o - v_o \exp(-2At) + x_o^2 \exp(-2At)] - K_2 y^2 \right] \\ &= K_1 y v_o - K_2 y^2 + \frac{K_1 y (x_o^2 - v_o)[1 - \exp(-2AT)]}{2AT} \end{aligned} \tag{37}$$

Then, with respect to this time-averaged expectation, y is optimized,

$$\frac{\partial \left\langle V_3' \mid x_o y \right\rangle_T}{\partial y} = \frac{\partial}{\partial y} \left[K_1 y v_o - K_2 y^2 + \frac{K_1 y (x_o^2 - v_o)[1 - \exp(-2AT)]}{2AT} \right]$$

$$= K_1 v_o - 2K_2 y + \frac{K_1 (x_o^2 - v_o)[1 - \exp(-2AT)]}{2AT} = 0 \tag{38}$$

$$y = \frac{K_1 (x_o^2 - v_o)[1 - \exp(-2AT)]}{4K_2 AT} + \frac{K_1 v_o}{2K_2} \tag{39}$$

Finally the $\left\langle V_3' \mid x_o y \right\rangle_T$ for this optimum y is expected over probable values of x_o, where $\{x_o\} = \{x\}$

$\left\langle V_3' \right\rangle_T = \int_x [\max_y \left\langle V_3' \mid x_o y \right\rangle_T] \{x\}$ finally results in

$$\frac{K_1^2 v_o^2}{4K_2} \left[1 + \frac{1}{2A^2T^2} - \frac{\exp(-2AT)}{A^2T^2} + \frac{\exp(-4AT)}{2A^2T^2} \right] \tag{40}$$

from which $\frac{C'}{T}$ is subtracted to yield $\langle V'_{net} \rangle_T$ as before in equation (17).

EXAMPLE FOR A CONVENTIONAL TRACKING CONTROL SYSTEM

Suppose for the same input as above $\langle V \mid xy \rangle$ is the negative squared "error" between input and output, $-(x-y)^2$, making the system a conventional tracking control system and making -V a "cost." The development in this case proceeds exactly as before. The corresponding results only are given below:

1. $\langle V_1 \rangle = -v_o$, where y is simply set to zero (41)

2. $\langle V_2 \rangle = 0$, where y follows x perfectly (42)

3. The information value is thus

$$\langle V_2 \rangle - \langle V_1 \rangle = v_o \tag{43}$$

4. If the supervisor samples x intermittently and optimizes y continuously,

$$\langle V_3 \mid t \rangle = -v_o[1- \exp(-2AT)] \tag{44}$$

5. The expected averaged return over a sampling interval, including the sampling cost, is

$$\langle V_{net} \rangle_T = -v_o[1 - \frac{1}{2AT} + \frac{1}{2AT} \exp(-2AT)] - \frac{C}{T} \tag{45}$$

6. If, instead of optimizing y continuously, the supervisor can only adjust y when he samples x, the result is

$$\langle V'_{net} \rangle_T = -v_o[1 - \frac{1}{A^2T^2} + \frac{2 \exp(-AT)}{A^2T^2} - \frac{\exp(-2AT)}{A^2T^2}] - \frac{C'}{T} \tag{46}$$

CONCLUDING REMARKS

A model has been presented to characterize the sampling behavior of an optimal supervisor of a deterministic, one-dimensional control process, where the value or return function based upon process input (and/or output) and control variable is available and well behaved. Future work may extend this model to multi-dimensional control problems and considerations of partially

random processes and value functions.

If there is a significant cost of updating y, the control setting or strategy, apart from the cost of sampling input x, then the supervisor may do better to sample x more often than he modifies y, changing y only if the new optimal y nets a difference in return which is greater than the cost for modification.

Sampling may be rendered not only in time, but also in degree, with more precise samples costing more. It would seem that the present theory could be enlarged in this direction.

A final word about determining the value return function $\langle V \mid xy \rangle$, upon which this whole paper is predicated. This is, and must always be, the fundamental and most difficult problem of human decision-making.

REFERENCES

[1] R.A. Howard, "Information value theory," IEEE Trans. on Systems Sciences and Cybernetics, Vol. 2, No. 1, 1966.

[2] N. Weiner, The Extrapolation, Interpolation, and Smoothing of Stationary Time Series, New York, Wiley, 1949.

[3] R.D. Smallwood, "Internal models and the human instrument monitor," IEEE Trans. Human Factors in Electronics, Vol. 8, No. 3, 1967.

[4] J.R. Carbonell, "A queuing model of many instrument visual sampling," IEEE Trans. Human Factors in Electronics, Vol. 7, No. 4, 1966.

[5] T.B. Sheridan, "Three models of preview control," IEEE Trans. Human Factors in Electronics, Vol. 7, No. 2, 1966.

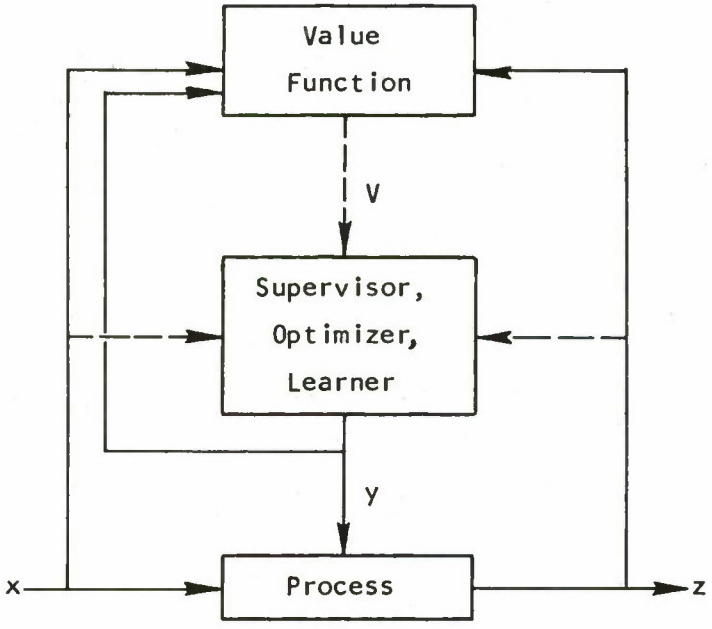


FIGURE 1: TYPICAL DIAGRAM OF OPTIMIZING OR LEARNING CONTROL SYSTEM

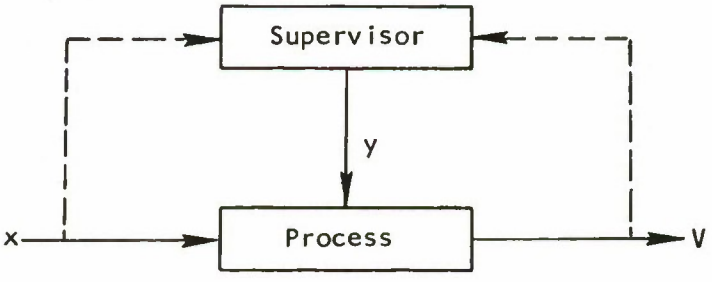


FIGURE 2: REDUCTION TO SYSTEM OF INTEREST TO THE SUPERVISOR

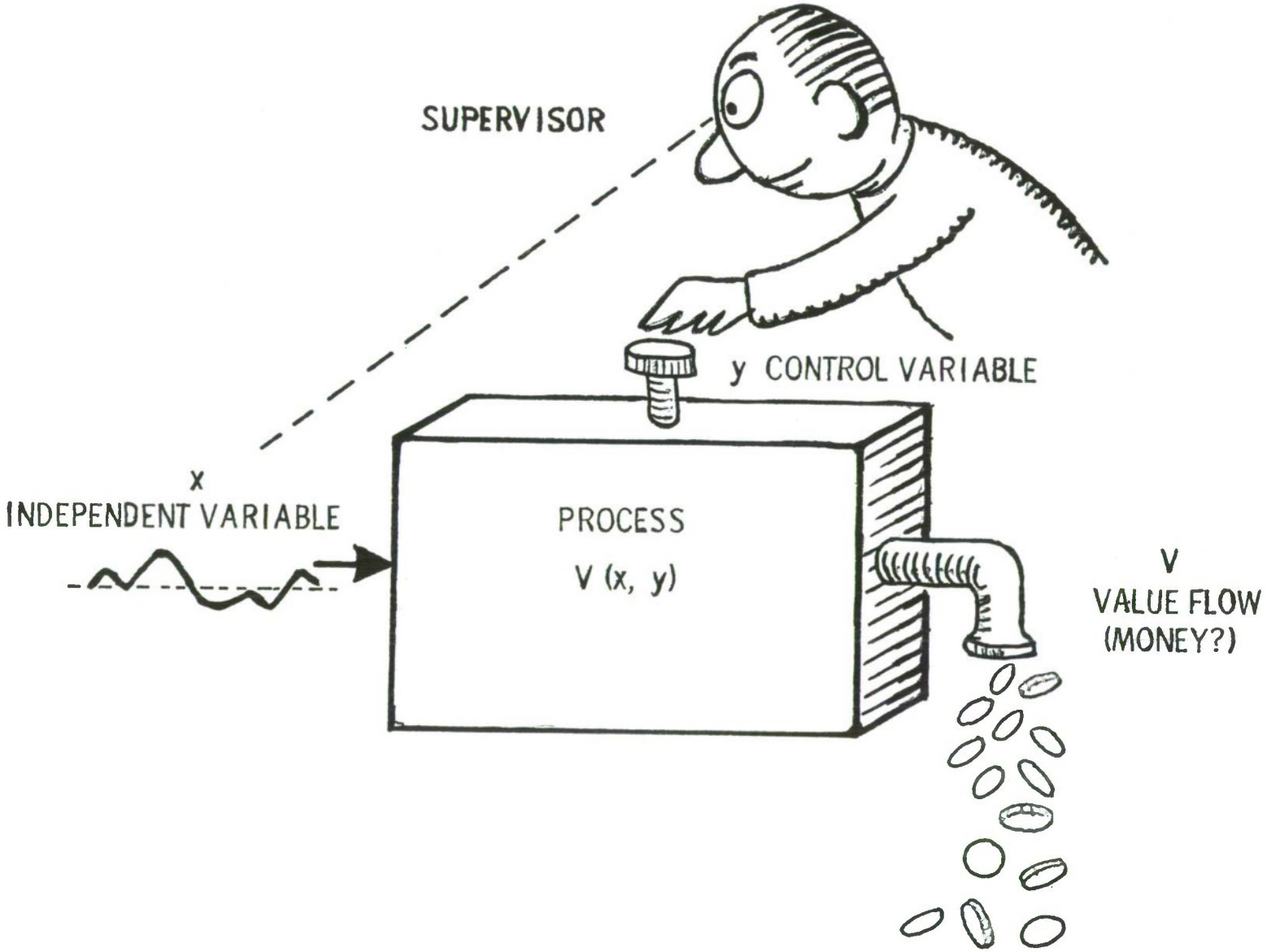


Figure 3: Caricature. The Supervisor Sampling Problem

GIVEN	DETERMINE
$\{x\}$, prior distribution of x	how often to sample x
V (x, y), value per unit time	how often to adjust y
C, cost of sampling x and/or adjusting y	what value of y to set in
$\{x' \mid x_0\ t\}$, a model of the supervisor's knowledge of x at some time t after sampling x_0	

Figure 4: The Formal Problem of the Superivsor

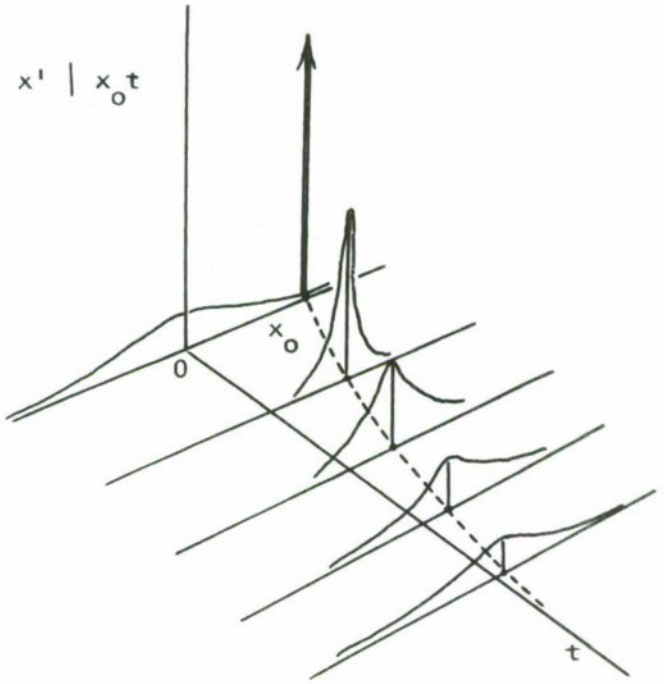


Figure 5: Variation of State of Knowledge x' With Time t After Sample x_o

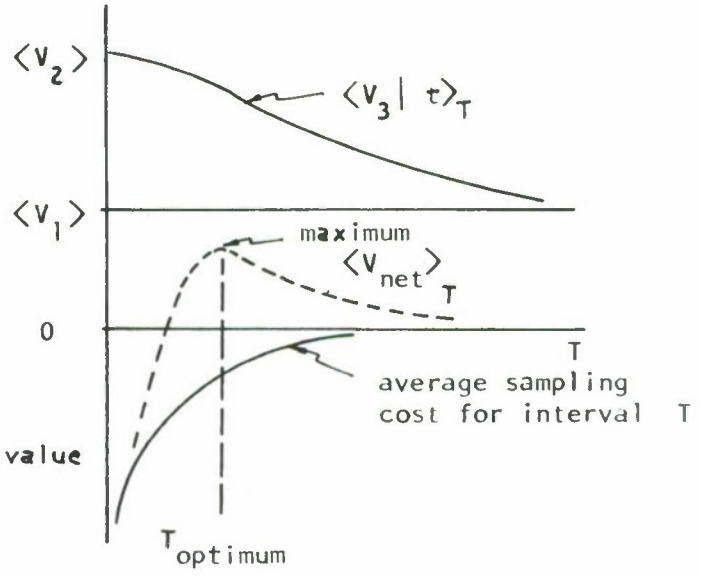


Figure 6: Determination of Optimum T

SYSTEMS ENGINEERING, SOCIAL SCIENCE, AND THE QUALITY OF URBAN LIFE

Ezra S. Krendel
Professor
University of Pennsylvania

Prepared for the Philadelphia City Planning Commission

February 18, 1970

Revised April 7, 1970

ABSTRACT

The purpose of this paper is to explore the applicability of the approach and the techniques of systems engineering to certain urban problems. Since there is ample evidence that systems engineering can be an effective tool in the design and operation of organizations to accomplish such goal oriented urban problems as: police scheduling, waste disposal, river purification, fire house location, and so forth, this paper addresses the relatively unploughed ground of applying systems engineering to the "quality of urban life". The quality of urban life may be implied by the extent to which a city manages to idenfity and to achieve the goals of its citizenry. These goals derive from a value structure held by this citizenry, and whether the generating values are constant or not, these goals may change as environment and technology serve to provide different expressions for values. Social indicators may be thought of as measures of the extent to which these goals have been achieved. A conceptual framework emphasizing an adaptive urban sub-system is presented, and in situ data are presented to illustrate the feasibility of the approach. Implications for planning and future work are included.

I would like to thank Ralph B. Hirsch of the Philadelphia City Planning Commission for his many suggestions and contributions throughout this study and John M. Greiner of the University of Pennsylvania for his work in analyzing the data; and Judith Milestone of the University of Pennsylvania for her work in gathering this data.

I. SYSTEMS

a. Concepts and Definition

There are many definitions of systems and systems engineering. I will use the following definition and enlarge on its implications in the ensuing pages.

> A system is a collection of interacting diverse functional units such as biological, human, machine information, and natural elements, integrated with an environment to achieve a common desired objective by manipulation and control of materials, information, energy, and life.*

The clear requirement for urban systems is to assess the extent to which "the common desired objective" has been achieved.

A city is an assembly of systems, some of which exist on a paralleled level with one another, and others of which are parts of a hierarchy of systems. These systems acting in concert generate the characteristics of the city. It is immediately clear that useful description of these sub-systems and their component processes is a necessary analytic procedure. Defining the interfaces between urban sub-systems , however, and specifying the interactions and redundant sub-systems is an equally important part of developing an effective understanding. A closed loop feedback structure for so doing was

* IEEE Group on Systems Sciences and Cybernetics, May 1967.

suggested by the author in a previous paper [Krendel, 1967]. In brief, a simple social ecological model was sketched out for which social well-being was suggested as one of the "common desired objectives." The functional characteristics were those common to any closed loop feedback system.

These are:

I. The first requirement is a definition of the desired set of goals and objectives, generally as a function of time, so expressed that the extent to which they were achieved could be reduced to measurable quantities. This functional quantity is the system command or forcing function, and for a regulator may be defined as a null command -- that is to say, "maintain the status quo".

II. A device or devices is necessary to sense discrepancies between the objectives or goals and the actual output of the system. This sensor or comparator is in many ways the heart of the system, because the accuracy of its assessments determine the feedback controls which the ecological system will impose on its environment or on its components in order to bring the measured discrepancy down to an acceptable level.

in the urban context, the citizenry is a major comparator or sensor, and the dimensions along which this comparison is carried out can be thought of as social indicators.

iii. The word "integrated" in the definition implies the third functional aspect. This is the controlling logic and system decision making which determines the duration and the extent of internal corrective responses which maintain the stability of the ecological system. In an urban system this process may be carried out by the political organizations, by the health and welfare agencies, by the police, and by the fire and the sanitation departments. Different problems, of course, require different agencies and demand different response times.

iv. For the system to function it must have an actuation or implementation capability. This is an urban equivalent of biological "muscle". This could be the police, a check sent through the mails, a telephone call by an official, and so forth.

v. Finally, there must be a means for tying the various components of the system together in a network which transmits information or energy. In a biological

system these would be the nervous system. In an urban system they are the many ways by which the components communicate -- by letter, by rumor, by word of mouth to a political figure, by violence, and so forth.

b. Adaptation

Systems which can be characterized by negative feedback loops have two characteristics which make them very attractive conceptual structures for describing biological or social systems. These characteristics are an ability to desensitize the system's output to external disturbances impinging on the system as well as to desensitize the system to variations in its internal components. These characteristics apply to the simplest sorts of linear closed loop feedback systems which maintain their system parameters unchanged and are restricted to linear operations on the signals which they pass. Thus, in a very broad sense [Truxal, 1961; Zadeh, 1963] they are adaptive systems, albeit passive adaptive systems. The active adaptive system is capable of changing its response characteristics as a function of environmental, component, or input signal changes which might be beyond the range of a linear feedback system to accomodate. Thus adaptation, as Truxal points out is more a point of view than a hard and fast system characteristic. It is the active adaptive system--the one which in effect reconfiguresitself when the environment changes too drastically -- which is most productive of insights into the

functioning of living systems. Such adaptation imposes two intuitively obvious requirements on the system: 1) it must be able to identify what is happening with enough detail to respond, and 2) it must be able to respond to the challenge by executing an appropriate repertory of actions. Identification may mean no more than recognizing heat from cold, and actuation may merely be moving away. What may be critical, however, is that both identification and actuation take time. They are not instantaneous.

Consider some very simple biological examples to fix the point: The retraction of the hand from a hot stove is an example of feedback system on the biological sub-system level. Walking out of the sun on a hot day and into the shade is an example of such behavior on a total system level. In both examples the control characteristics of the system, i.e., the musculature, the heat sensors, and so forth do not change as the system responds. If we can conceive of and write the underlying equations we would say that their parameters were fixed. These are examples of passive adaptive behavior in the general sense. In social systems and in urban systems we would expect the more complicated form of active adaptive behavior in which components of the system change so as to maintain a desired level of behavior and control. To continue the biological example, such species that were not able to adapt in time have, in general, become extinct. The dinosaurs, for example, were not capable of changing their mode of

living so as to accommodate changes in the environment. Similarly the koala bear which lives on a particular form of eucalyptus shoot is not a candidate for a long, and flourishing existence on Earth. Man and his institutions, on the other hand, are prototypes of the successful adaptive processes.

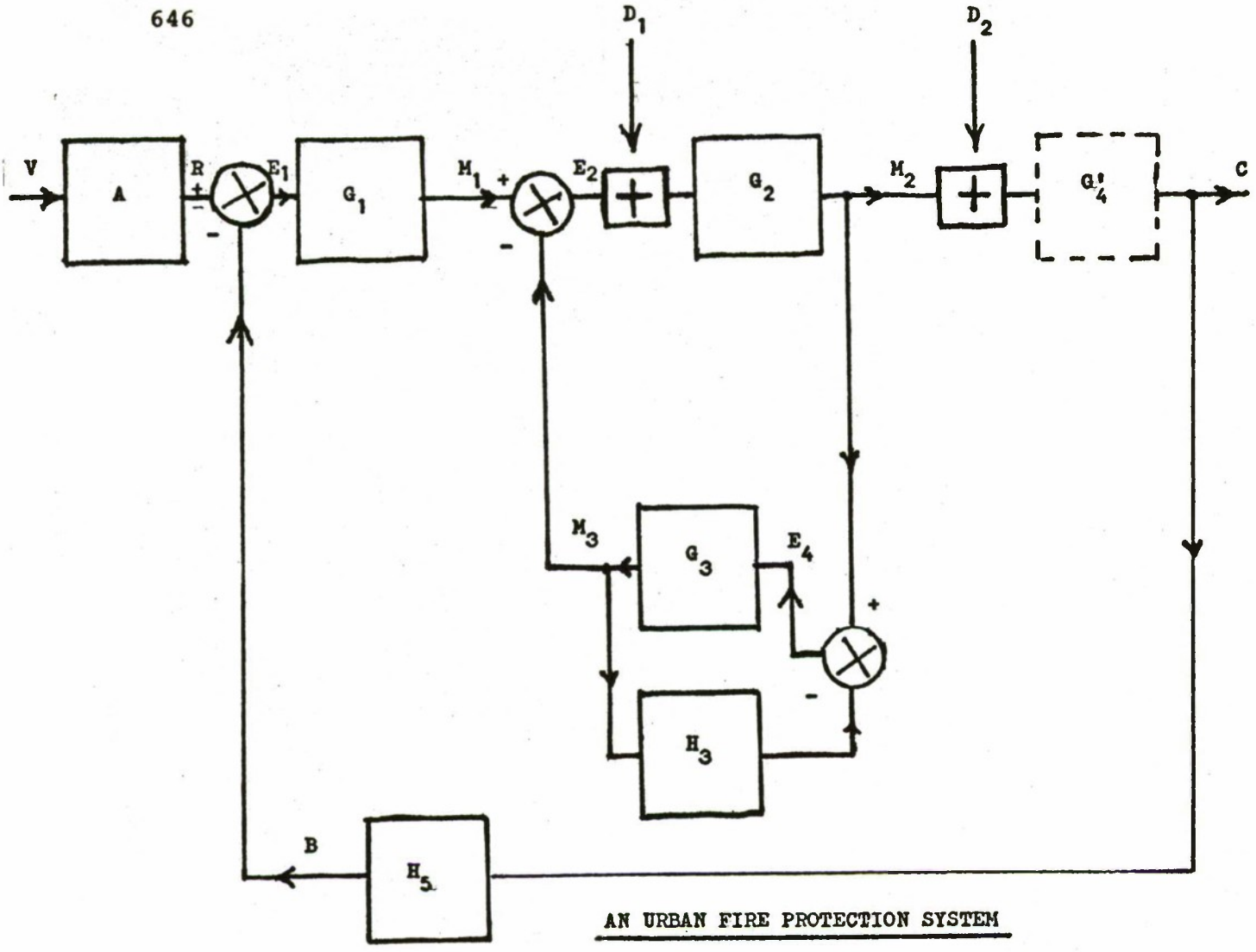
Adaptive behavior in the biological or the social context has hitherto required an evolutionary time scale to manifest itself. Peruvian Indians, Eskimos, etc., provide examples where metabolism and respiration, and the associated body processes have actively adapted and evolved to meet the requirements of a hostile environment. In the social context there exists a literature and a fairly convincing set of arguments about the stabilizing and regulating influences of ritual and tradition in the life of primitive societies. In particular, in island cultures where the resources are limited, tradition and ritual are devices which help to maintain the ecological balance. [Rappaport, 1968]. In more advanced societies where resources are limited, custom may dictate such corrective responses as late marriage, becoming a military mercenary, or entering a celibate clergy. When we examine urban problems the urgency and time scale is such that we cannot wait for solutions based on ritual or custom to evolve; we must sense their need and determine how we can encourage their development. Urban dynamics and adaptive systems are discussed in this presentation to suggest meaningful solutions in an urban

context which may make the accomplishment of specified goals more likely, and within a meaningful time interval.

c. A tutorial example: Fire protection

There are many interacting sub-systems which, in concert, create and maintain our urban societies. Police, Health, Fire departments, Education, Transportation, and so forth, all readily come to mind as organizationally defined sub-systems. By examining a representative sub-system the importance of characterizing process components will become clear. In addition alternative emphases within a sub-system will be made explicit, so that the implication of policy decision can be seen. In general, for urban sub-systems with strong or direct implications for the quality of life, the measures of system performance are difficult to establish. Our selected example will be chosen so that the performance measure is easily defined so as to focus attention on sub-systems components and interactions. For expository purposes consider a highly simplified functional block diagram of a City Fire Protective System. The Fire Department itself, it will be seen, is part of the system. Even in this contrived and primitive case it can be seen that the conventional organizational subdivisions in a city may not be appropriate to the current problems. The analogies between a Fire Protective System and such other functions of the city as perhaps: Traffic Safety, Health, or Economic Opportunity Creation, and many other socially oriented services are obvious.

Unfortunately the detailed exposition of the analogical similarities into functional block diagram form is somewhat more elusive. In Figure 1 a functional block diagram for a City Fire Protection System is presented. Conventional control theory notation is used [Dorf, 1967] in that upper case letters represent transformed variables. They are manipulated as if they were conventional Fourier Transforms, and since neither stability conditions nor time domain responses are calculated, this use of block diagram algebra does not require the variables be specified as Fourier, Laplace or perhaps Z transforms. It is similarly unimportant at this time whether the configuration in Figure 1, or an elaborated version, be examined by paper and pencil analysis or by a combination of analysis and digital simulation. Although there is strong support for the digital simulation of systems which may have hard non-linear behavior [Forrester, 1969], there is a well developed counter position supporting the imaginative use of analysis [Graham, 1969]. At this stage a simulation might tend to obscure rather than to reveal the concepts involved in the analysis. Each of the blocks in the diagram represents a process which converts an input to an output. Each process is characterized in some sense by a magnitude (how big a response?) and by a time relationship with the input or an accepted reference (How long did it take to respond or did it anticipate?). Some of the processes are administrative or legal and some are mathematical such as determining a probability of a fire occuring. The following definitions apply:



AN URBAN FIRE PROTECTION SYSTEM

- R - Typically: # of fires of greater than \$X damage/time interval/area
- G_1 - Transforms E_1, system performance discrepancy, to M_1, # of building units inspected for fire hazards/time interval/area
- G_2 - Transforms E_2, # of units/time/area failing inspection to M_2, the probability of a fire starting/time/area
- G_3' - The closed loop representation of the inspection process, $\frac{H_3}{1 + G_3 H_3}$, whose output, M_3, is # of units passing inspection/time/area
- D_1 - Introduces the effects of malfeasance of inspectors
- D_2 - Introduces fires started by arson or by events unrelated to inspection
- G_4' - Closed loop description of fire department actions in fighting fires
- C - System output: #of fires of greater than \$X damage

Figure 1

V --- is the primary input to the fire protection system and as such is a measurable physical variable. Although specific interests may dictate other directly measurable variables; a reasonable and convenient choice would be: number of fires/time/area. The area selected could be a census tract, an economically homogeneous neighborhood, a major subdivision of the city and so forth. The time interval could vary from weeks to years.

A--- is the transformation operator which converts the primary input to the controlled variable. The value of this transformation could very well be unity .

R ---, the controlled variable, must be considered together with the performance criterion for the system. Thus if R were:

a. number of fires of greater than $X damages/time/area
b. number of persons killed or injured in fires/time/area
c. number of persons made homeless/time/area.

In all of these and other similar ones we clearly want the performance criterion to be less than or equal to some selected number relating to one of the criteria above.

A simplifying point of view would be to consider the fire protection system to be a regulator against fires, and to define R identically equal to zero. In this case the control system in Figure 1 would attempt to drive the error, E_1, to zero. Since achieving the goal of $E_1=0$ is not feasible, a cost and a benefit would have to be

assigned to various degrees of proximity to this goal for meaningful evaluation.

G_1--- transforms E to the number of buiding units inspected per time interval per area, which is M_1. The rules for this inspection, which is for the purpose of detecting and eliminating potential generating sources for fires, or conditions which might stimulate existing fires, are modifiable on the basis of the indicated feedback in the subsequent block which describes the inspection process.

G_2 --- This transformation converts E_2, the number of units/time/area failing inspection to the probability of a fire starting, M_2

D_1 ---This is a disturbance input which adds to the number of units/time/area failing inspection by introducing the effects of malfeasance of inspectors as well as chance errors.

$$G_3' = \frac{G_3}{1 + G_3H_3}$$

is the closed loop representation of the inspection process whose output is the number of units passing inspection/time/area. G_3 describes the lags, cycle time and effectiveness of the inspection, and H_3 represents the penalties imposed on the owners of property which fails to meet requirements. These penalties are in general fines and have long lag times associated with them.

D_2 --- This is the disturbance input which characterizes the starting of fires by arson or by events not related to an inspection procedure.

G_4' --- Contains the closed loop function which describes the actions of the fire department in attempting to contain the fire and to limit the damage. The G_4' block represents the sensing decision making and command processes, and its feedbacks are the plant dynamics of the response time of the men and equipment and their effectiveness in action. The time lags associated with this loop are order of minutes rather than the days and weeks, and perhaps even months or years for the previous blocks.

C ---- is the system output, and is either directly commensurable with R, or the error criterion, or can be made so by the action of H_5. In the particular example shown, C is expressed as the number of fires of greater than $X damage, amd H_5 is unity.

The foregoing, however abstracted and simplified, is illustrative of many of the issues involved in the applications of systems techniques to the various urban systems. The first issue, of course, is where to draw the conceptual boundary about the "system". One could relate fire protection services to insurance rates and to similar influences on the business and industrial climate of the city. It is clear that a rapid escalation in complexity could occur. The choice of the boundary is based on both pragmatic and intuitive considerations. Pragmatic because a system model or configuration is of little or no value if its component parts and their interactions cannot be reduced to some form

of quantitative measurement. Intuitive because at the problem formulation stage of a study the analyst is free to express and to exercise what he feels is comfortable and tractable choice of boundaries between his interests and external "chaos". This choice is clearly highly personal, and interacts with the pragmatic considerations. For example the fire department loop with its various components can be described in gross summary statistics for response and effectiveness or it can be explored in great detail so that the "nerves and brain" of the system are dissected and the location of fire stations, types of equipments, methods of operation and so forth are examined.

Many potential inputs have been omitted from Figure 1. The inspection process, G_3, is influenced by: the budget constraints on either municipal or contractual employees, the skill level of the labor market, the morale of the organization performing this task, the interests of casualty insurance companies as well as of owners and bankers, and last but not least the political influence of the residents. The feedback term H_3 contains within it the enforcement structure and the penalties associated with failure to comply. Such penalties can range from warnings and fines to condemnation of buildings. Clearly the enforcement block will be characterized by the same types of delays which plague our court system. The foregoing augmented list of inputs or considerations are still by no means complete. They serve to indicate the role that educated judgement must play in

formulating the system configuration for the base case. Thereafter iterations will successively reveal either the need for embellished structure, or the analyst may be able to discard components of little or no effect on the variables of interest. This outcome of the analysis is of patent importance in both the political and the planning process.

Proceeding from Figure 1, we can, after a little algebra, determine an expression for the output of this simplified City Fire Protection System. The output, C, is in terms of the number of fires/unit time/area of greater than $X damage:

$$C = \frac{D_1 G_2 G_3' G_4'}{1 + G_3' G_2 + G} + \frac{D_2 G_4' (1 + G_3' G_2)}{1 + G_3' G_2 + G}$$

where $G = G_1 G_2 G_3' G_4'$

It is not clear without data and component dynamics whether a fire protective system within the restricted definition of this system in the foregoing discussion could be improved most cost effectively by improving the inspection procedure or by improving the fire department's response capability. Analytically, and within this restricted context, there is no difference how one proceeds. There are obvious implications, however, such as: the deleterious effects on urban life of fire engines racing about the streets, the implication

that the city will not intervene for the citizen short of disaster, and so forth. The difference resides in values, not analysis; and these values imply and help to define the quality of life in the city. The existence of values imposes on the city, which has the goal of a high quality of life, the need to adapt in an unusual manner. In common usage an adaptive system is one which can maintain its response or output within desired limits despite changes in the environment, the system's components, or the driving function. Adaptation to a citizen value structure implies the capability to reexamine the selected performance criteria and to alter them on the basis of value oriented inputs. An example can be found in the opposition to high raise housing units and preference for single or small multiple dwelling units, as an expression of changing values .

From a highly oversimplified point of view, the political structure assimilates the citizens values, and creates a set of goals which the values, in some concensus, imply. Should these goals be related to social well being or life quality, then social indicators are needed to determine the extent to which these goals have been achieved. Thus, in this context, the urban dwellers'values are primary since they help to determine the goals to which the indicators point, and in fact may change these goals. If such a change in goals is likely, the measurement of values must be carried out in an ongoing manner to enable the making of real time decisions rather than to provide historical reference points.

The city should establish the institutional structure which can measure and respond in time to indicated change. A sensing capacity without an actuation or response capability is meaningless. The city must be prepared to create new and effective responses, and to reconfigure its processes as the circumstances demand.

In the next section some representative values of process dynamics for a selected administration structure, as well as an expression of elicited values will be presented.

2. MEASUREMENTS AND DATA

a. In Situ Dynamic Measurements

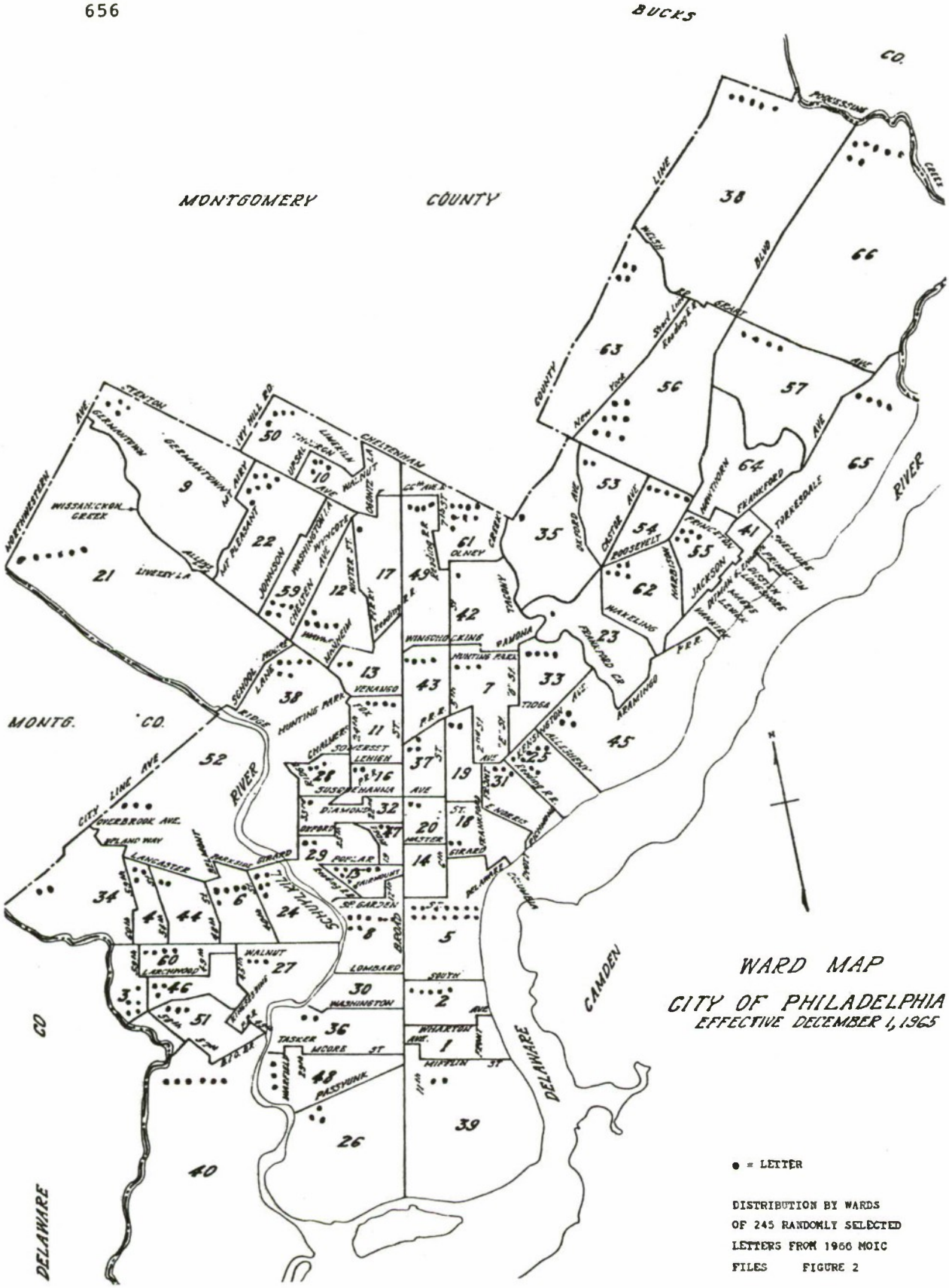
As a general rule it is difficult to justify the measurement of dynamic characteristics for a human dominated system by open loop techniques. To be realistic and representative the measurement must be made during the normal course of the operation of the system. Thus if we were measuring the delay times in a postal service, we would not use a tracer letter which was identifiable by the operators of the system. We would not want them to stop everything to hurry the test missive along its route. We would want the operators to be burdened with their usual responsibilities to other letters. When we examine a feedback system within the urban context our data requirements are different and more complex than a simple speed of service measurement. We are interested in both the closed loop response and component behavior. Consider the fire department block in Figure 1. An appropriate input/output closed loop measurement for this system would be the time from the sending of the alarm to the first direct action against the fire taken by members of the department. Enclosed within this closed loop response time are: dispatch and message center times, the time to get the equipment and men rolling, transit time, and set up and command decision times on site. These measurements are straightforward since it is clear and in fact compelling when a fire alarm is sounded, and the response is equally straightforward. What of responses

which relate to "quality of life"? Here we can scarcely identify the input, let alone assign a unique response to it.

There is a possible way by which inputs related in some sense to life quality can be uniquely related to the city responses . This can be accomplished by using the files of the Mayor's Office of Information and Complaints, MOIC. Although the time resolution is in days, and although the sample is biased toward the literate and the somewhat knowledgeable, and although this is not the only source of freely generated citizen complaints in the city (L & I, Police, etc.), it provides a valuable laboratory case. MOIC is a unique test facility for two reasons: 1) By their questions and complaints the citizens reveal their own values as well as their estimate of the extent to which the city has achieved goals which satisfy these values, and 2) The unique linking of letters from citizens to the assigned municipal response enables us to follow an input to its closed loop resolution.

In order to get a rough check on the geographic representativeness of the MOIC files a map of the City of Philadelphia wards, with the distribution of 245 letters processed by MOIC in 1966 was prepared, see Figure 2. There appears to be an underrepresentation in the North Philadelphia wards.

To understand the actual form of the original data and the limitations which this imposed upon its subsequent analysis, consider the procedure by which the citizens' letters to MOIC are handled. On


BUCKS CO.
MONTGOMERY COUNTY
MONTG. CO.
DELAWARE CO.
RIVER
DELAWARE RIVER
CAMDEN
WARD MAP
CITY OF PHILADELPHIA
EFFECTIVE DECEMBER 1, 1965
● = LETTER
DISTRIBUTION BY WARDS
OF 245 RANDOMLY SELECTED
LETTERS FROM 1966 MOIC
FILES FIGURE 2

first writing to the Mayor's Office, a citizen is assigned a file in which copies of all subsequent relevant information (letters to the complainant, referral actions and dates, memos to the mayor himself concerning the matters, etc.) are retained. The same file is used for all subsequent correspondence by the given individual, whether or not it concerns the original problem. The only exception to the rule of one individual complainant per file is the case of multiple letters on a single narrow subject (such as the closing of a particular bar, or the appointment of a city official); these are placed in a single file. The filing system was instituted about 1965 and the available records covered the period 1964 to 1968 at the time the sample was taken.

If possible, the citizen's inquiry is handled entirely within MOIC (e.g., acknowledgement, rejection of crank letters, etc.) and records are made of that office's action, the relevant dates, and the contents of the letter. If MOIC itself cannot dispose of the communication, the letter is transmitted to an appropriate department. If multiple complaints are received in a single letter, copies of the letter are transmitted simultaneously to all appropriate departments, the referral date and target departments being noted. The complainant is also informed by MOIC of this transmittal. Notices of any actions by the departments (letters or visits to the complainant, memos to the mayor,etc.) are forwarded to MOIC and filed under the complainant's I.D. number. If a department re-routes the letter once more, the dates and

actions of the departments to which this second referral is made are also obtained and recorded by the Mayor's Office. Thus, the file under an individual's I.D. number in the Mayor's Office consists of a history of the progress of each matter noted by that individual in his letter to MOIC, complete with the dates of receipt, referral, and final disposition of each matter and copies of all associated internal and external communications occasioned by the processing of the matter.

Two samples of the files were obtained in the summer of 1968. The first was a trial study of 50 files from 1964 and 1966, the second of 250 files from 1966. The contents of every hundreth file were summarized, acting under the plausible assumption that since the files were arranged by the chronological order of the receipt of the initial letter, the existing arrangement represented a random sequence of inquiries. It was necessary to go through the entire set of the available files several times, following the sampling procedure, to obtain a total of 300 files. The detailed data in the files were condensed into a summary format which included all relevant dates, the department involved, and an outline of the complaints and any subsequent replies from the departments. An illustration is presented in Table I in the following pages. The data were also punched on cards. In addition to the basic information on dates, number of complaints, department affected, and the location of the complaints (by ward and census tract), the coded information includes inferences concerning the ultimate disposition of each complaint, the area affected by the complaint, the urgency and

TABLE 1
Illustrative Abstracted Raw Data
12 MOIC Files

Book Data	File Data	Actual Complaint	Ward	Census Tract
1. #2425 8/7- 8/7 PHA	initial letter- Mayor-C: 8/7	wants to be moved to ground floor of housing project due to sick child	2	3B
2. #2493 9/1 9/21 Police	i.l.rec'd: 8/27 M-C. 9/1 Pol.HQ. to M. 9/21	gallery had been broken into; wants better protection from police foot patrol	8	9B
3. #2778 9/1 9/21 Streets	i. comp. in early July to Streets Sts-C. 7/10 2nd l. 8/17 Sts. to M. 9/1 M.-C. 9/21	Wants light in an alleyway so it won't be a gathering place for drunks	3	46D
4. #2456 9/1 9/15 Welfare	i.l 8/20 M-C 9/1 refd welfare 9/1 W.-M. 9/15	Has just had twins; is only 19 yrs. old; not eligible for aid due to short residence in city; where can she get help?	3	46E
5. #2822 10/16 10/30 County Ct.	i.l. no date M-C 10/16 refdCt. 10/16 Ct.-C. 10/30 action completed	wants more money from the Court for divorce action	Hatboro, Penna.	
6. #2709 9/17 10/15 L and I	i.l. not dated (recounts compls. of 2 wks. earlier to pol. and fire) M-C. 9/17 refdL&I 9/17 L&I-M 10/1 M-C 10/15 2nd l. 10/18 M.-C. 10/20 refd Sts. 10/20 Sts-C. 11/1 3rd letter 9/19/66 M.-C. 9/27 refd L&I 9/27 L&I-C. 10/13 4th l. 10/7	complaints about dirty lots; death of husband; ethnic background of neighbors; that Negroes are cab drivers; that police need more power etc.	20	20E

TABLE 1 Cont.

7.	#2470A 8/18 9/21 L&I	i.l 8/12 M-C 8/18 refd L&I 8/18 L&I-M. 9/4 M.-C 9/21 2nd l. 10/7 action incomp.	water in the cellar	15	15C
8.	#2424 7/22 8/21 Police L & I	i.l. not dated recd M. 7/16 M-C 7/22 refd L&I and Police 7/22 reply to HQ from dist. 7/30 HQ-M. 7/31 L&I-M. 8/21 Police compl. action; L & I did not.	abandoned truck with refrigerator inside.	62	55C
9.	#3550 11/13 12/2 Welfare	i.l. 11/10 M-C 11/13 refd W. 11/13 W.-M 12/2 action in process	needs a refrigerator in which to keep food etc.	43	43F
10.	#4052 10/3 10/6 Police	no i.l. C. interviewed by Pol. 10/2 Dist-HQ 10/3 HQ-M 10/6 case closed	unruly teenagers	22	22K
11.	#2774 9/24 10/6 Police	no i.l. refd 9/24 answer from Magistrate 10/5 (incl. hearing report from 8/20) C. Int. 9/30 dist-HQ 10/1 HQ-M 10/6 case closed	behavior of police while ticketing; wanted investigation	New Jersey	
12	#2489 8/19 9/21 Police	i.l 8/3 (refers to 2 earlier letters sent to L&I M-C 8/19 refdP. 8/18 Pol.-M 9/16 action taken Pol.-M. 9/21	unlicensed dog	25	25A

KEY TO 300 FILES

1-25 1964
26-250 1966

Bd of Ed	Board of Education
Bonnelly	same as County Cort
camp. supp.	campaign support
cc.	carbon copy - usually means Mayor did not get first copy of i.l.
C	Census tract
City rep.	City representative
Cty Bd Asst	County Board of Assistance
CPC	City Planning Commission
comm.	commendation
D.A.	District Atty.
Dis-Hq(Q)	police district report to commissioner (often just send xerox of above to Mayor)
D.R.P.A.	Delaware River Port Authority
FPC	Fairmount Park Commission
Hq-M	final police report to Mayor
i.l.	initial letter
L&I	Licenses and Inspections
Mng Dir	Managing Director (also Corleto)
Mun Guides	Municipal guides
MUC	Manpower Utilization Commission
M-C	Mayor reply to Complainant
MOIC	Mayor's office of Information and Complaints
PAAC	Phila. Anti-poverty Action Committee
P.O.	Post Office
PHA	Phila. Housing Authority
pkg.	parking
Pub Info	Public Information
P. or Pol.	police
P.P.	Public property
pers.	City Personnel
Rec	Dept. of Recreation
R.D.A.	Redevelopment Authority
Sts.	Streets
(t)	refers to thank you letter for action done (or thanks)
Traf tick	traffic ticket
Vets Admin	Veterans Administration
We	Welfare (or Welf.)
Wa	Water
W	Ward

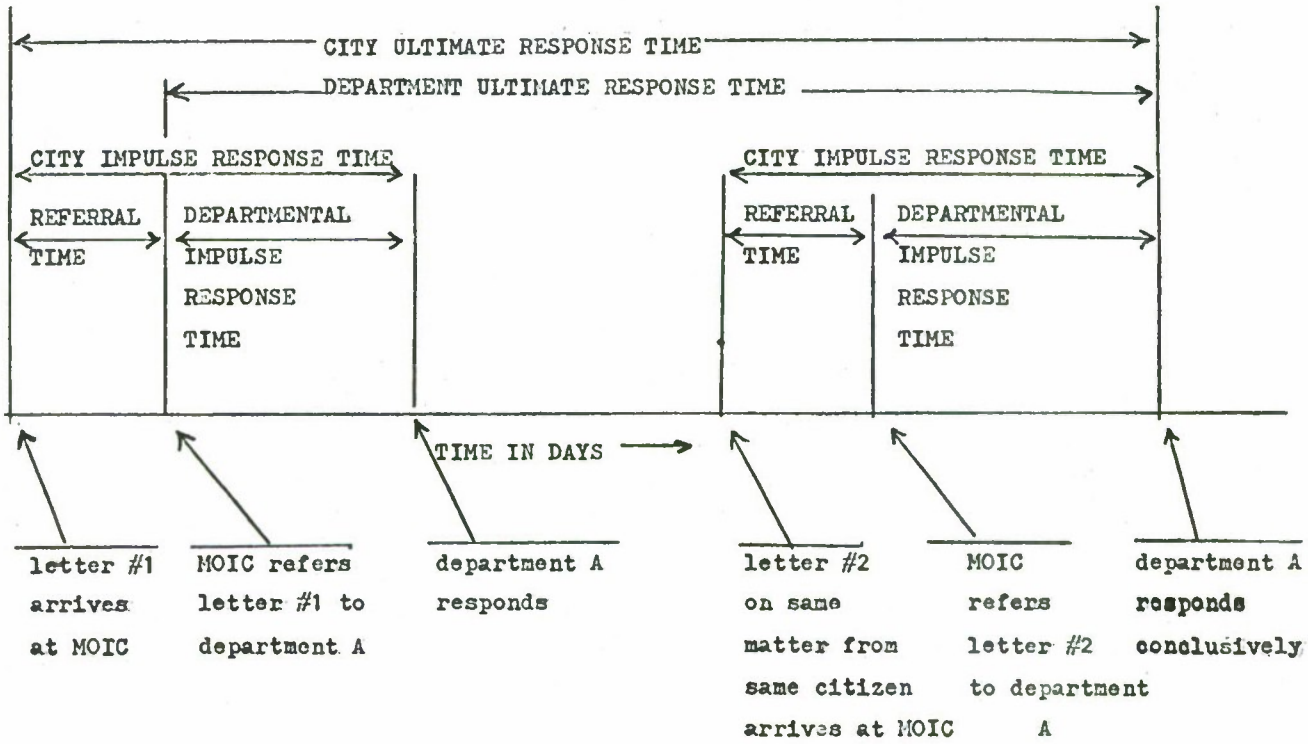
reasonableness of each matter, and several classification of the actual concept of the complaint.

These two forms of the 300 file sample from MOIC - the condensed data summary and the encoded data - constitute the basis for the analyses reported in this paper.

b. Analysis of the Data

There are two types of analysis which were conducted. The first measured the process dynamics of the city agencies when confronted with a citizen request. Since the purpose of this analysis was to obtain time constants characterizing MOIC as well as the final action department, it is natural to examine these time durations from the lights of a control engineer. Figure 3 illustrates the nature of the measurements. Impulse response time is the time to the first significant response, and ultimate response time is the time to solution for the system. The citizen writing the complaints is conscious only of the ultimate city response time, as indicated in Figure 3. The administrative structure underlying this interval is of little interest to him, but it is of interest to a systems analyst since it provides the data from which process dynamics can be constructed.

The second type of data derives from the content of the letters as from the departmental assignments made by MOIC. From these data, unstructured



STRUCTURE OF RESPONSE TIME MEASUREMENTS FIGURE 3

though they may be, the underlying dimensions of "quality of life" indicators may be detected.

1. Time response data

Composite distributions for city impulse response time, ultimate response time and a partition , based on letter content, of impulse time into information and into action responses are presented in Figures 4,5, and 6. These curves indicate that on the average, the citizen must wait 23 days for the first response to his letter and 36 days for a conclusive response. There is a 90% probability that the first response will arrive within 42 days and that the matter at hand will be settled within 54 days. The distributions of low response times in Figures 4,5,6 were inflated because of a data gathering convention. The effect is that letters arriving at MOIC in 0 or 1 day were bunched with letters arriving two days after being posted. It was found that importance of the letter as indicated by content did not seem to have any effect on its speed of processing. Emergency letters or calls, of course, would not go through this office. A possible comparison is with NYC, [Savas, 1969] where the claim is made that the Mayor's Action Center, which is a round the clock operation handling 200,000 phone calls and 5,000 letters per year, guarantees a ten day response.

Distributions such as Figure 4 may, because of their physical source as well as their shape, be expected to follow models arising from queuing or servicing time. The Erlang or gamma type distribution is an immediate candidate.

This distribution which is

$$f(t) = \frac{\lambda^{k}}{(k-1)!}\, t^{k-1}\, e^{-\lambda t} \; ; \; t \geq 0$$

has the following characteristics

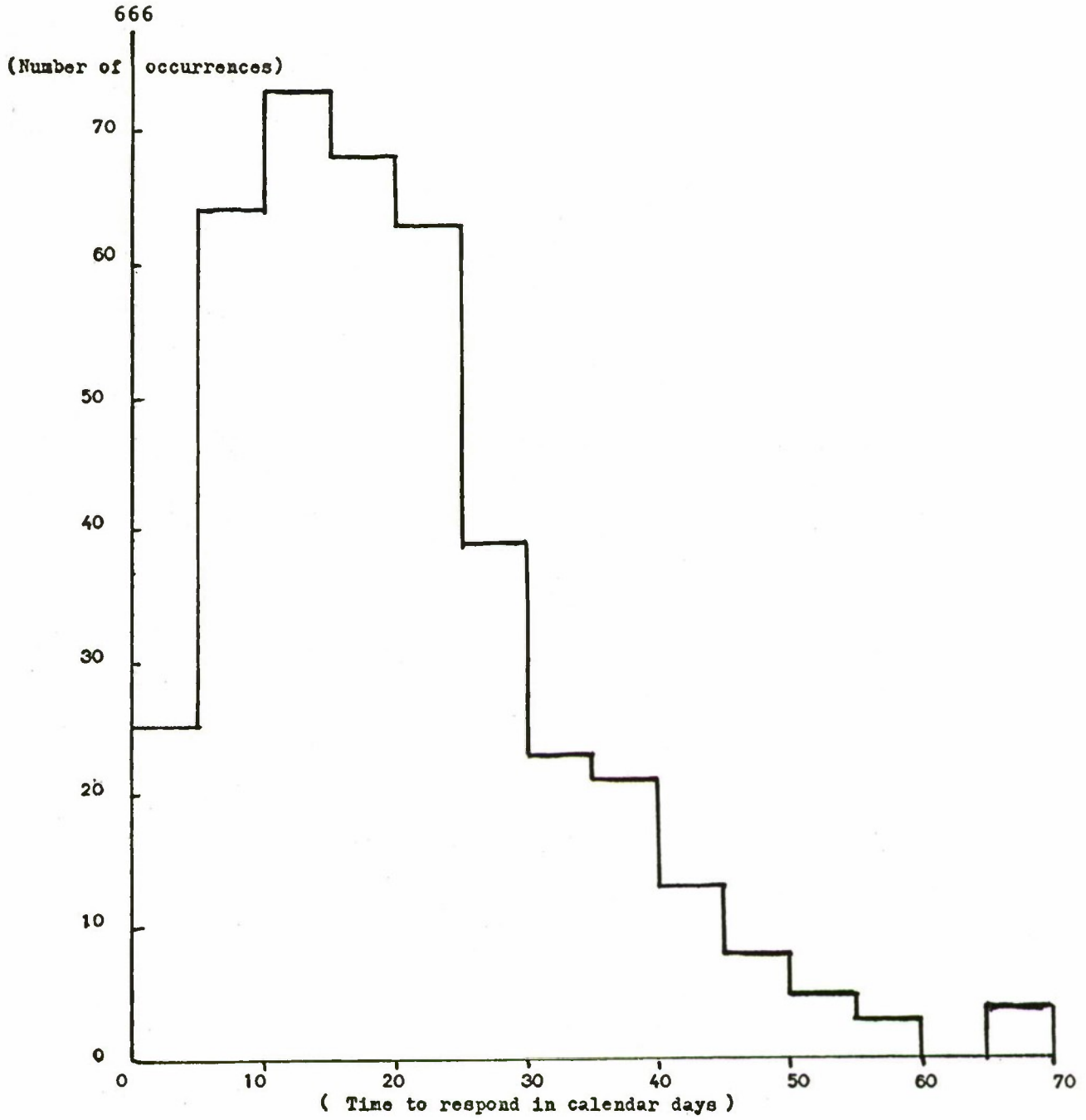
mean, $\bar{t}$, $= k/\lambda$

standard deviation, σ , $= \sqrt{k}/\lambda$

mode, m, $= (k-1)/\lambda$

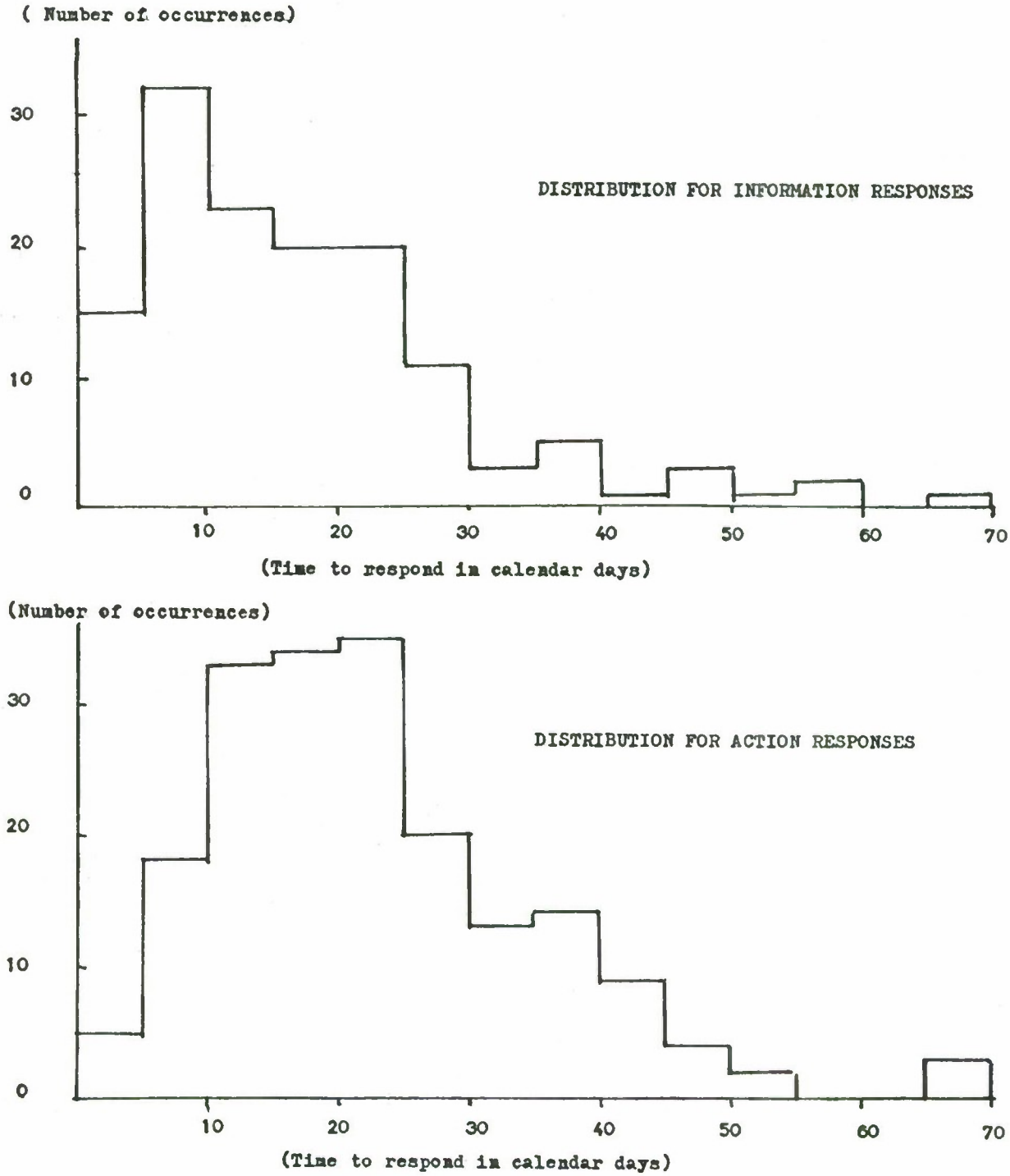
[Greiner, 1969] demonstrated an adequate fit for k=2 for the distribution on Figure 4, such that the determinations for $1/\lambda$ from mean, standard deviation and mode were: 11.4, 12.5, and 15.4 respectively. An average value of $13.1 = 1/\lambda$ was used for the fitted distribution.

Similar fits of varying degrees of closeness were made to other distributions such as those on Figures 5,6, and 7. In order to convert these time functions to the operator notation of Figure 1, Laplace transforms were taken, thus generating the second order transfer functions in Figure 8. These functions are indicative of a potentially fruitful procedure, rather than being convincing in and of themselves. If we can specify administrative process dynamics, we can by understanding the demands or the process, specify operational characteristics which the component functions must exhibit for the criterion behavior to exist. In other words, we can suggest changes in configurations so that the city can be effective in sensing and responding. We can do this by analogy to Figure 1, if we can set performance criteria, define sensor metrics,



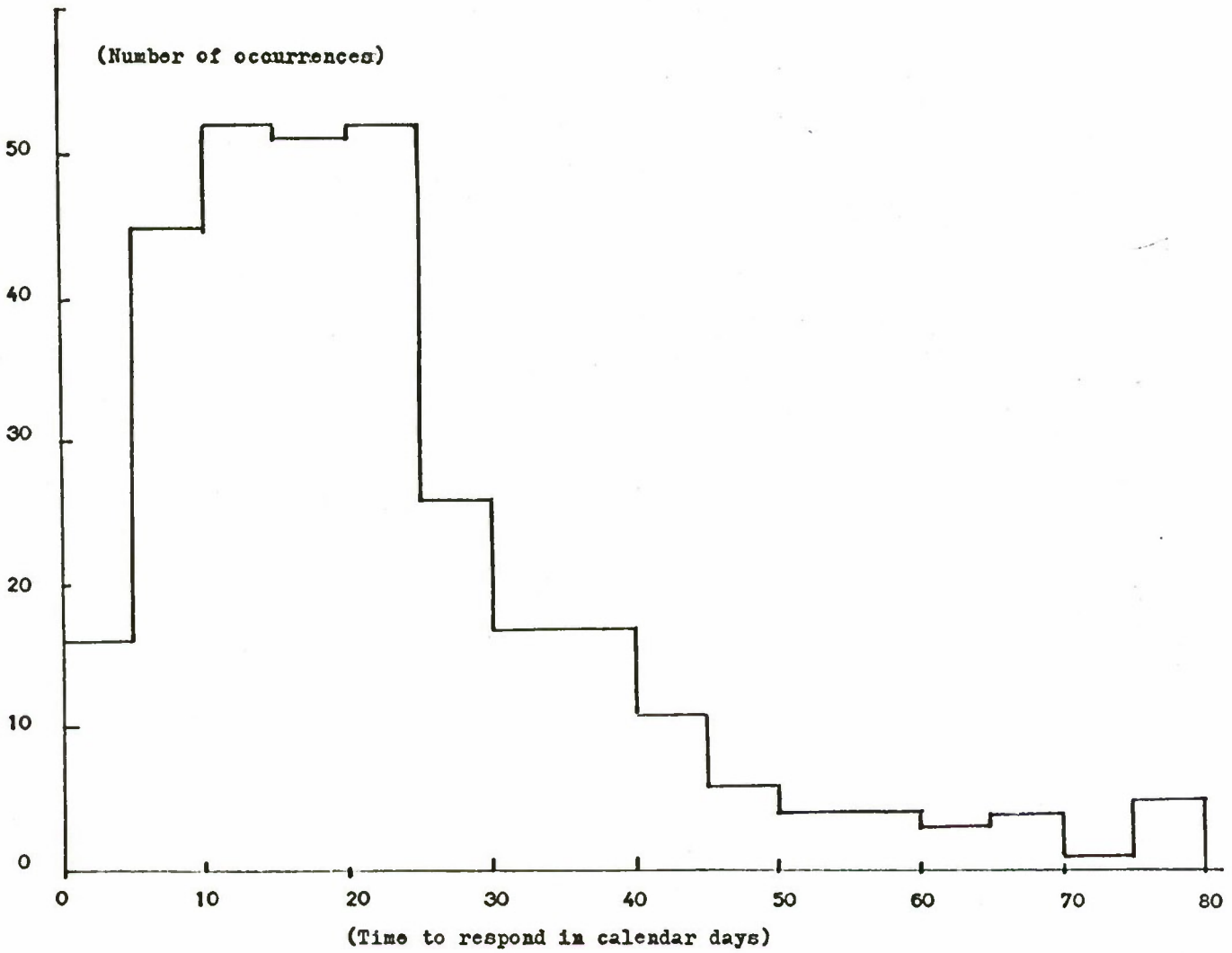
COMPOSITE DISTRIBUTION OF CITY GOVERNMENT "IMPULSE" RESPONSE TIMES

Figure 4



PARTITION OF CITY GOVERNMENT "IMPULSE" RESPONSE TIMES INTO INFORMATION & ACTION

Figure 5



COMPOSITE DISTRIBUTION OF TIMES FOR "ULTIMATE" CITY GOVERNMENT RESPONSES

Figure 6

and measure component dynamics.

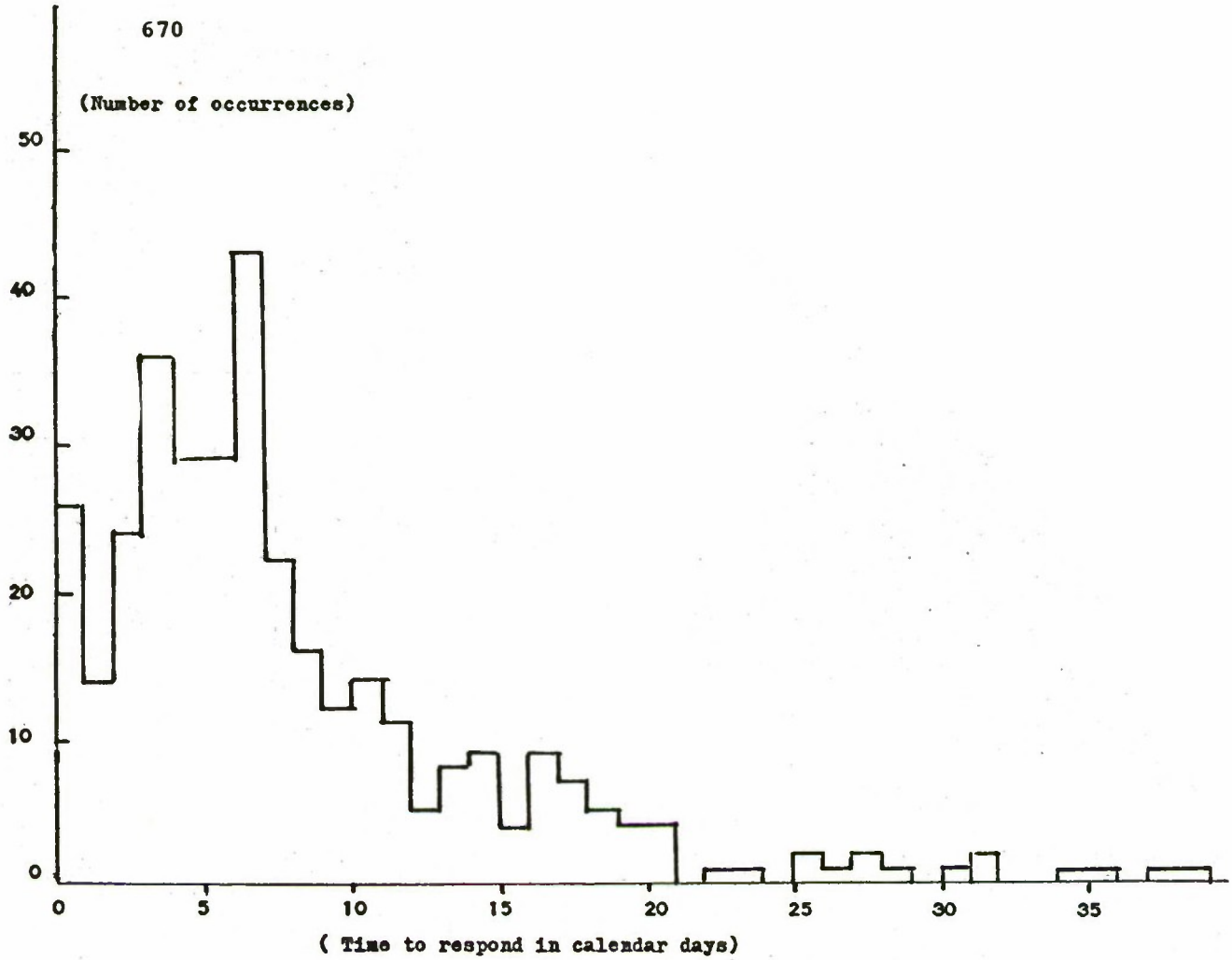
2. Content associated data

The nature of the unstructured data as can be seen in Table i is such that sophisticated efforts at scaling these data would be in vain at this time. The data can be classified and ranked to generate insights into the values of the citizenry and into the appropriate "quality of life" indicators which such values imply.

A first cut at the content of the letters to MOIC can be obtained from the following distribution of the 300 files among the various city departments

DEPARTMENT	INITIAL DISTRIBUTION OF ASSIGNMENTS	ULTIMATE DISTRIBUTION OF ASSIGNMENTS
Police	27.6%	27.3%
Streets	20.1%	20.3%
MOIC	17.2%	15.7%
Licenses and Inspection.	12.0%	13.1%
Others	23.1%	23.5%

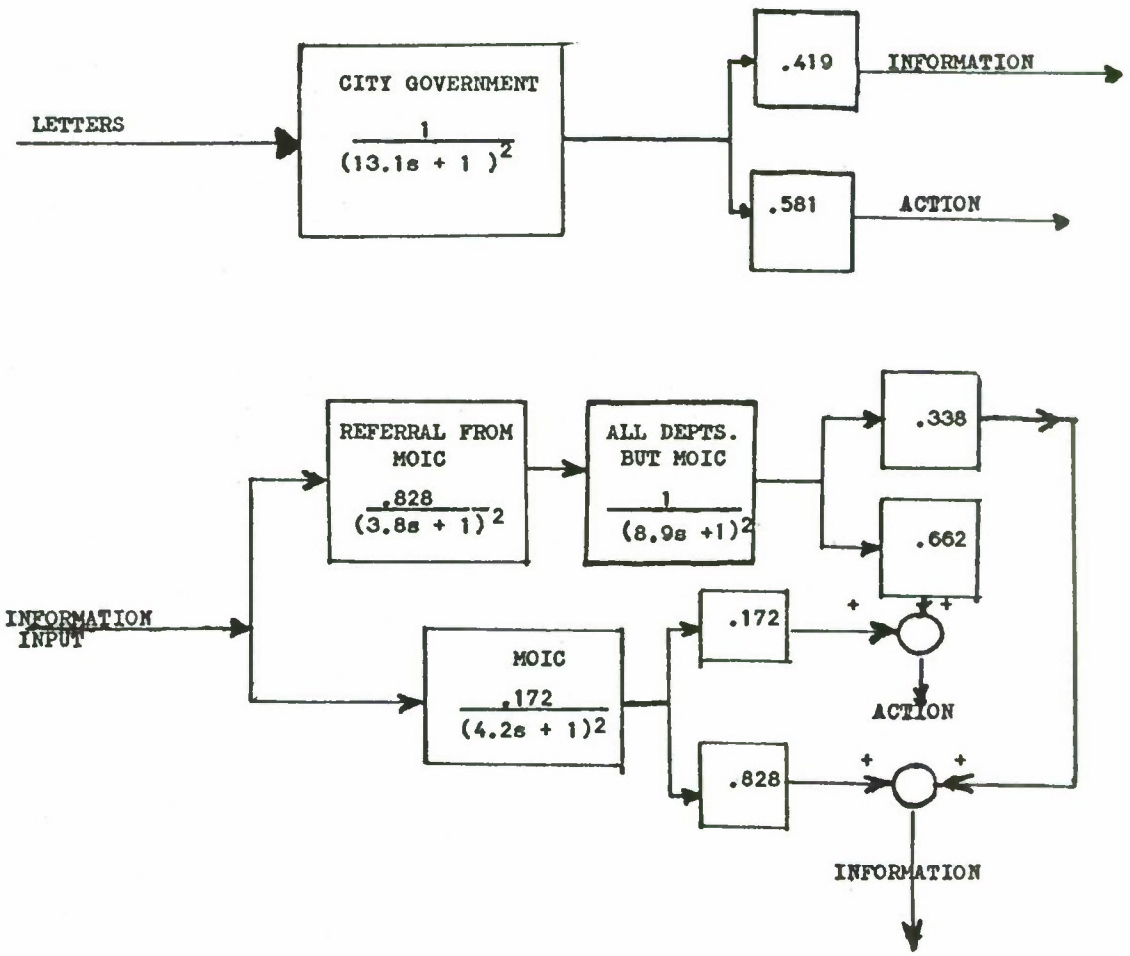
The small amount of variation between the initial and ultimate assignments indicates the validity and accuracy of the initial MOIC judgments.



COMPOSITE DISTRIBUTION OF MAYOR'S OFFICE OF INFORMATION AND COMPLAINTS , MOIC,

REFERRAL DELAY TIMES

Figure 7



PROCESS DYNAMICS DERIVED FROM IMPULSE RESPONSES

N.B. 1. Gamma probability density distribution of order 2 was fitted to data.
2. All times are in calendar days.

Figure 8

In a second effort to summarize the content of the files, approximately 400 "complaints" in the files were categorized in the following eight categories. (Some letters or files contained more than one complaint,hence a number greater than 300.)

Personal Safety	32%
Resource Requirements	13%
Information	16%
Personal Image	11%
Movement	11%
Health and Sanitation	10%
Housing and Shelter	5%
Culture and Recreation	2%

The foregoing listing provides a first cut based on MOIC information only at the ranking of interests and concerns of the citizenry which relate to the quality of urban life. Obviously, other data sources, both within the city structure and deliberately created as well, are needed to arrive at a more representative basis for designing "quality of life" indicators. Scaling these indicators, combining them into aggregate measures, and designing the data handling, the decision and the response systems is an exciting goal for the future.

3. CONCLUSIONS AND RECOMMENDATIONS

This study has demonstrated the potential value of looking at urban syb-systems, which bear directly on the quality of life in the city, as if they were feedback systems. It has been shown that this approach is feasible, but before it can be applied over a sufficiently large set of sub-systems to be effective, a larger, future study is required. This approach requires social indicators or quality of life indicators which are different from the conventionally presented indices [U.S. Dept. of H.E.W., 1969; Bauer, 1966]. The indicators needed must be capable of rapid response and of being used for real time decisions. These hypothetical indicators are in effect predictors, and if they were vaiid, their integrated value would be closely related to the more stable and sluggish conventional measures. These indicators should derive from the goal structure of the citizenry [Terleckyj, 1969], and be responsive to changes in this structure.

A group of recommendations for trial implementation which derive from systems engineering consideration are presented:

- By detailed systems engineering methods, exploit the possibilities for analyzing urban sub-systems so that the choices between alternative feedback paths will be made expiicit to the citizenry, technician and administration.

- Establish the urban system equivalent of test signals whose purpose would be to exercise the system at some irregular schedule so that current assessments of system performance could be obtained. A concept similar to the teiegrapher's "speed of service" criterion might arise from such a procedure.

- As a research target create mechanisms which become institutional forms for expressing opinion -- more frequently than at the conventional election cycle -- and without the violent release which characterizes some current protest.

- Develop procedures for combining information obtained from the census and by the police, welfare and other agencies with information from MOIC, and

- establish upward channels which may or may not be filtered and brought into perspective by local experts so that a better understanding of changing values and goals and perceived needs can be achieved.

- Create a cadre of trained observers for limited periods of service who are from the communities and capable of observing, detecting, and of filtering information so as to emphasize trends as well as communicating their intelligence to appropriate agencies.

- Establish communications channels to provide downward information about social services, employment opportunities, housing, social activities, and so forth, to respond to the information needs revealed by this study.

LIST OF REFERENCES

Bauer, R.A. "Social Indicators" The MIT Press, Cambridge, Mass., 1966

Dorf, R.C. "Modern Control Systems" Addison-Wesley, Reading, Mass., 1967.

Forrester, J.W. "Urban Systems Dynamics" MIT Press, Cambridge, Mass. 1969.

Graham, D. "Toward an Analytical Industrial Dynamics: A Production Distribution Problem Revisited" Management Science Center, University of Pennsylvania, 1969.

Greiner, J.M. "The Application of State Space Control Concepts to Urban Systems" term paper for OR 616, University of Pennsylvania, 1969.

Krendel, E.S. "A Simple Viewpoint on Urban Modeling" OECD Experts Meeting, London, 1967.

Rappaport, R.A. "Pigs for the Ancestors", Yale University Press, New Haven, 1967.

Savas, E.S. "City Halls and Cybernetics" in Cybernetics and the Management of Large Systems, ed. by E.M. Dewan, Spartan Books, Washington, 1969.

Terleckyj, N.E. "Measuring Progress Toward Social Goals: Some Possibilities at National and Local Levels" paper read at AAAS, Boston 12/27/69.

Truxal, J.G. "The Concept of Automatic Control" in Adaptive Control Systems, ed. by Mishkin and Braun, McGraw-Hill, N.Y., 1961.

U.S. Depart. of H.E.W. "Toward a Social Report", Washington, 1969.

Zadeh, L.A. "On the Definition of Adaptivity" Proc. IEEE, 51, pp.469-470, 1963.

THE EFFECTS OF CHANGES IN INPUT POWER SPECTRA ON HUMAN OPERATOR COMPENSATORY TRACKING

John C. Heifferon and Russell A. Hannen

Air Force Institute of Technology
Wright-Patterson Air Force Base, Ohio 45433

Abstract

An investigation is made into the effect on mean-squared error, as a performance measure in a single-axis compensatory tracking task, due to changes in the shape of the input power spectral density. Quasi-linear describing function theory is used to predict normalized mean-squared error with the pilot-vehicle dynamics represented by the crossover model. A digital computer program is used to evaluate the resultant mean-squared integrals and to plot normalized mean-squared error as a function of input and crossover model parameters.

The effect on human operator performance due to an added low-amplitude high frequency extension to a near-rectangular input spectrum is studied. The model predicts an optimal pilot gain in the mean-squared error sense when this high frequency shelf is included in the input spectrum. In the absence of the high frequency shelf, the locus of minimum mean-squared error is shifted to the stability limit and the model predicts a change in the human operator's behavior. In this case an optimum pilot gain would result in neutral stability. The human operator will, in the absence of the shelf, decrease his mean-squared tracking error to a point where a minimum positive phase margin is reached which is consistent with good closed-loop low frequency response. This application of the model to published experimental results lends support to the model predictions.

I. Introduction

Background

Analytical representations of biological processes have revolutionized man's understanding of his environment. In particular, mathematical models of the physiological subsystems of man himself have provided the capability to extrapolate to novel situations and to predict responses to environmental changes. In the design of manned flight vehicles, the requirement to match the vehicular characteristics to the characteristics of the human pilot has provided the challenge for the development of a satisfactory model of the human pilot subsystem. Over the past two decades there has been a considerable amount of effort extended in this pursuit.

The most useful pilot models to date have developed from classical servo-analysis of random-input quasi-linear systems. These models were developed from a single-axis compensatory tracking task defined by the block diagram of Fig. 1.

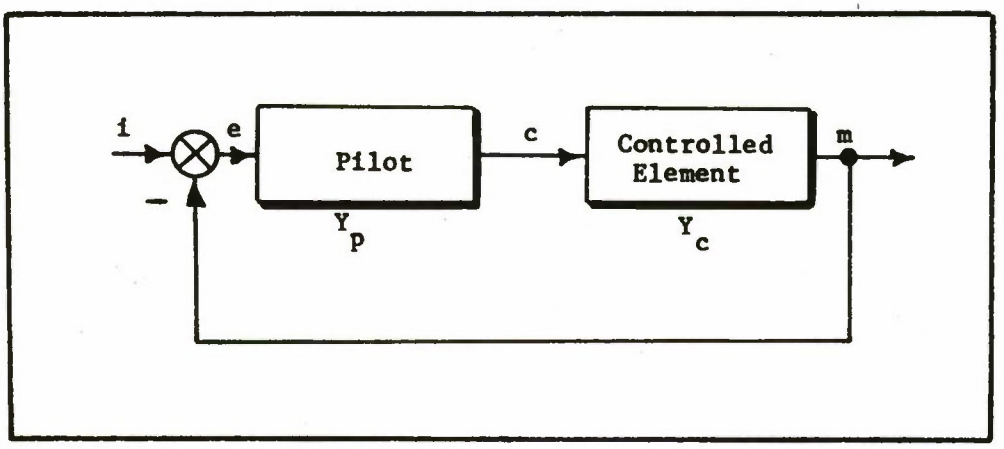


Fig. 1. Single-Axis Compensatory Tracking

A stationary random-appearing forcing function, i(t), is applied which results in a displayed error, e(t). The pilot output, c(t), acts on the fixed controlled element dynamics, Y_c, to produce the system output, m(t). The pilot characteristics in this task have been extensively studied and, with the aid of spectral analysis, a verbal-analytical pilot model has developed. The basic pilot model consists of (1) a generalized describing function consisting of a pilot gain, various lead and lag terms, and a time delay, and (2) a series of adjustment rules which estimate the range and choice of describing function parameters depending upon the input forcing function characteristics and the type of controlled element.

The Describing Function Form

The low frequency describing function form used in this investigation is (Ref 10:8)

$$Y_p = K_p \frac{(T_L j\omega+1)}{(T_I j\omega+1)} e^{-j\omega(\tau+T_N)} \qquad (1)$$

where Y_p is the pilot model transfer function. The pilot gain is K_p and the equalization characteristic is given by $(T_L j\omega+1)/(T_I j\omega+1)$ where T_L is a lead time constant and T_I is a lag time constant. The reaction time delay term, $e^{-j\omega(\tau+T_N)}$, encompasses an affarent term, τ, and a neuromuscular lag, T_N.

The pure time delay represented by $e^{-j\omega\tau}$ is due to the delay encountered in the central nervous system between sensor excitation and efferent output response. Tau has been observed to range between 0.1 sec and 0.2 sec and is taken to be constant with an individual.

The neuromuscular delay, T_N, decreases monotonically with increasing forcing function bandwidth. Observed variations in T_N have ranged from above 0.6 sec to below 0.1 sec.

The pilot gain, K_p, and the equalization characteristic, $(T_L j\omega+1)/(T_I j\omega+1)$, are the main elements used by the pilot in adaptation. The form of equalization inserted is generally compatible with a good low frequency closed-loop response and with system stability. The equalization forms selected for the three simple controlled elements of K_c, $K_c/j\omega$, and $K_c/(j\omega)^2$ are shown in Table I.

Table I
Pilot Describing Function Forms

Controlled Element Transfer Function Y_c	Equalization Adjustments	
	Low Frequency ($\omega << \omega_c$)	Mid-Frequency (ω about ω_c)
K_c	$\frac{K_p}{(T_I j\omega+1)}$; $\frac{1}{T_I} << \omega_c$	$K_p(T_L j\omega+1)$ $\frac{1}{T_L} > \omega_c$
$\frac{K_c}{j\omega}$	K_p	$K_p(T_L j\omega+1)$ $\frac{1}{T_L} > \omega_c$
$\frac{K_c}{(j\omega)^2}$	$K_p(T_L j\omega+1)$ $\frac{1}{T_L} << \omega_c$	K_p

The Adjustment Rules

In addition to the adaptation of the describing function form to achieve stability and good closed-loop response, a trained operator is known to follow an internal optimalization criteria which sets the describing function parameters in a way which is thought to be analogous to that of minimizing mean-squared tracking error.

The Crossover Model

It has been demonstrated in detail that the open-loop function, Y_pY_c, is well approximated for compensatory tracking by the frequency response function of the two parameter crossover model

$$Y_pY_c(j\omega) = \frac{\omega_c e^{-j\omega\tau_e}}{j\omega} \qquad (2)$$

The effective reaction time delay, τ_e, and the gain crossover frequency, ω_c, are both adjustable parameters in this model. The adjustment selections for the controlled elements K_c, $K_c/j\omega$, and $K_c/(j\omega)^2$ are listed in Table II (Ref 10:28).

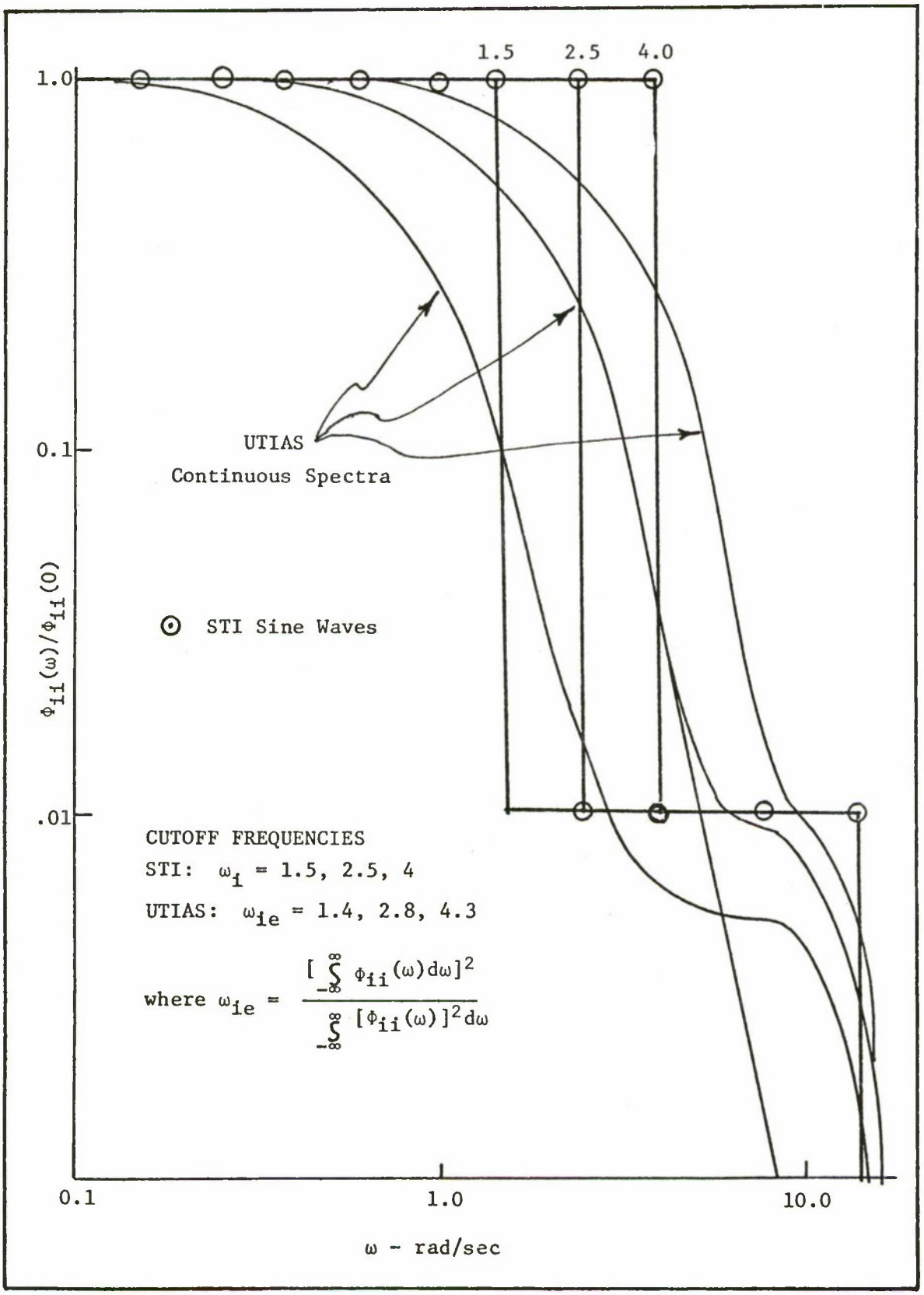


Fig. 2. STI and UTIAS Power Spectra

Table II

Crossover Model Describing Function Forms

Controlled Element Transfer Function Y_c	Operators Describing Function Y_p	Crossover Model (ω about ω_c)
K_c	$\frac{K_p e^{-j\omega\tau_e}}{(T_I j+1)}; \frac{1}{T_I} << \omega_c$	$\frac{K_p K_c}{T_I} \frac{e^{-j\omega(\tau+T_N-T_L)}}{j\omega}$
$\frac{K_c}{J\omega}$	$K_p e^{-j\omega\tau_e}$	$K_p K_c \frac{e^{-j\omega(\tau+T_N-T_L)}}{j\omega}$
$\frac{K_c}{(j\omega)^2}$	$K_p e^{-j\omega\tau_e}(T_L j\omega+1); \frac{1}{T_L} << \omega_c$	$K_p K_c T_L \frac{e^{-j\omega(\tau+T_N)}}{j\omega}$

Experimental Studies

A comprehensive study of compensatory tracking was completed by Systems Technology, Inc. (STI) (Ref. 10) in 1965. The STI forcing functions, i(t), consisted of a sum of ten sine waves of random phasing and logarithmic spacing. The STI input power spectral density thus consisted of a sum of ten discrete delta functions. The three cutoff frequencies were defined by setting equal amplitudes for the lowest six, seven, or eight frequencies and by attenuating the amplitudes of the remaining frequencies by 20 db. The high frequency segments extending beyond the cutoff frequency were intended to allow human response measurements in the critical gain crossover region. These low amplitude segments were considered to not materially affect the operator's low frequency performance.

More recent studies (Ref 4:13) were conducted at the University of Toronto Institute of Aerospace Studies (UTIAS). The UTIAS input forcing functions were similar in nominal shape and bandwidth to the STI input forcing functions but were derived from the filtered output of a Gaussian noise generator. The UTIAS forcing functions thus had a continuous power spectral density.

Fig. 2 compares the continuous UTIAS spectra with the discrete spectra of STI. The extended low amplitude segments of the UTIAS spectra are seen to be attenuated a nominal 20 db and to correspond to the high frequency portion of the STI spectra. This attenuated portion of the input spectral shape is commonly referred to as a high frequency shelf. The middle frequency UTIAS power spectrum where ω_{ie} = 2.8 rad/sec is shown with and without the high frequency shelf.

In the UTIAS studies a comparison was made between the aggregate averaged mean-squared errors obtained by trained subjects with and without the high frequency shelf. The comparison revealed that the shelf had a marked effect on operator performance. Removal of the shelf caused a large reduction in score and increased the amplitude ratio with a consequent increase in crossover frequency. There was no change in phase. This represented a simple

increase in operator gain upon removal of the high frequency shelf. This result did not agree with the STI assuption that a high frequency shelf attenuated by 20 db would not appreciably affect an operator's performance.

An Analytical Approach

The approach used in this investigation can be explained with reference to the block diagram of Fig. 3. The block diagram depicts the crossover model for a single-axis compensatory tracking task being disturbed by a frequency limited white noise input process. The control loop signals are represented by their power spectral densities, e.g., $\Phi_{ee}(\omega)$.

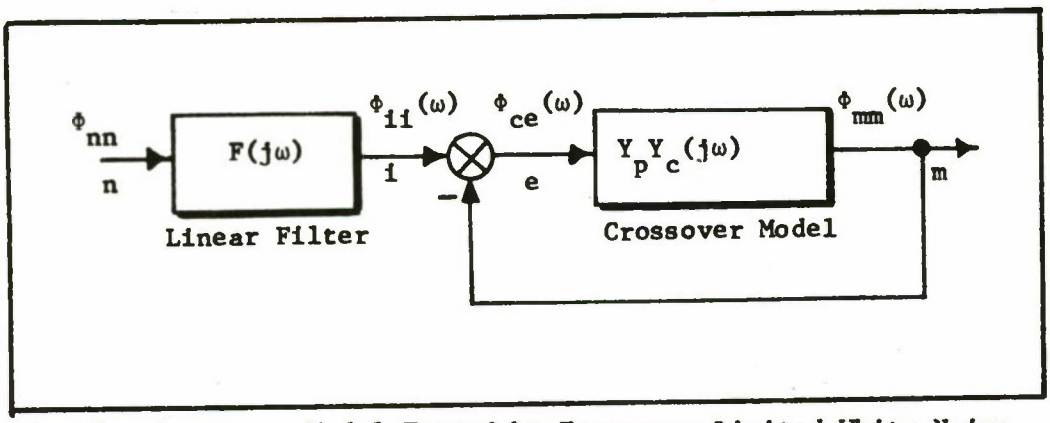


Fig. 3. Crossover Model Forced by Frequency Limited White Noise

The difficulty of handling an exponential term in closed-loop analysis led to the use of a first order Padé approximation for the time delay factor of the crossover model

$$e^{-j\omega\tau_e} \doteq -\frac{(j\omega-2/\tau_e)}{(j\omega+2/\tau_e)} \tag{3}$$

The frequency response function representing the pilot/vehicle dynamics for the crossover model then becomes

$$Y_pY_c(j\omega) \doteq -\frac{\omega_c(j\omega-a)}{j\omega(j\omega+a)} \tag{4}$$

where $a = 2/\tau_e$

The n(t) process, besides possessing a zero mean Gaussian amplitude probability distribution is assumed to have a power spectral density of white noise. White noise is the nonrealizable time signal which has a constant power spectral density from zero to infinite frequencies. The linear filter, $F(j\omega)$, is designed to shape the frequency spectrum of the white noise process to approximate the spectral shape of the input being studied. The relationship between the power spectral densities of the input and the output of the linear filter is

$$\Phi_{ii}(\omega) = |F(j\omega)|^2 \Phi_{nn} \tag{5}$$

where Φ_{nn} = constant power spectral density of white noise.

The power spectral density of the error signal, e(t), is given by

$$\Phi_{ee}(\omega) = \left|\frac{E(j\omega)}{N(j\omega)}\right|^2 \Phi_{nn} \tag{6}$$

The mean-squared value for the error signal, e(t), is defined to be

$$\sigma_e^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} \Phi_{ee}(\omega)\, d\omega \tag{7}$$

Substituting Eq. (6) into Eq. (7) we have

$$\sigma_e^2 = \Phi_{nn}\frac{1}{2\pi}\int_{-\infty}^{\infty} \left|\frac{E(j\omega)}{N(j\omega)}\right|^2 d\omega \tag{8}$$

Similarly, the variance of the input forcing function, i(t), is given by

$$\sigma_i^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} \Phi_{ii}(\omega)\, d\omega \tag{9}$$

Substituting Eq. (5) into Eq. (9) we have

$$\sigma_i^2 = \Phi_{nn}\frac{1}{2\pi}\int_{-\infty}^{\infty} \left|F(j\omega)\right|^2 d\omega \tag{10}$$

The integrals in Eq. (8) and (10) are of the general form of

$$I_n = \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} \left| \frac{C_1 s^{n-1} + C_2 s^{n-2} + \quad + C_n}{A_1 s^n + A_2 s^{n-1} + \ldots + A_{n+1}} \right|^2 ds \tag{11}$$

and have been tabulated by Phillips and others (Ref 5:259). A digital computer subroutine (Ref 2) is used in this investigation for the evaluation of the above integral. A listing of this subroutine is included in Ref 6.

II. Problem Development

It was shown in the investigation of Ref 10 that, with the pilot/vehicle dynamics represented by the two parameter crossover model, a rectangular input spectrum would predict a minimum mean-squared error component due to the forcing function at the model stability limit. That is, minimum tracking error would occur at the zero phase margin point. The results of this rectangular input forcing of the crossover model are illustrated in Fig. 3. The figure represents normalized mean-squared error versus gain crossover frequency with forcing function bandwidth as a parameter. The crossover frequency and input bandwidth are normalized by τ_e.

It is demonstrated (Ref 10: 58) that in the region of interested the normalized mean squared error obeys the "one-third law" given by

$$\sigma_e^2 = \frac{1}{3}\left[\frac{\omega_i}{\omega_c}\right]^2 \tag{12}$$

as long as the ratio ω_i/ω_c is less than one. As a result of the STI experimental results with discrete input spectra it was further stated that the human operator approximates the phase margin adjustment rule given by

$$\phi_M = \frac{1}{3}\left[\frac{\omega_i}{\omega_c}\right] \tag{13}$$

again when the ratio ω_i/ω_c is less than one.

Now, if a human operator approximates the phase margin adjustment rule of Eq. (13) in a system which obeys the "one-third law" represented in Fig. 3 then he cannot be minimizing mean-squared error. Minimum mean-squared error for this model occurs at the point of zero phase margin. Conversely, if a pilot is operating with a positive phase margin and is minimizing mean-squared error, then his operating point cannot be described by both the phase margin adjustment rule of Eq. (13) and the "one-third law" model.

Anderson (Ref 1) demonstrated that the predicted minimum mean-squared error occurred with a phase margin which was representative of experimental results for some continuous input cases. In particular, Fig. 4 illustrates the results for the crossover model forced by white noise filtered through a simple second-order lag filter given by

$$F(s) = \frac{K}{(s+\omega_b)^2} \tag{14}$$

The locus of minimum mean-squared error shown by the dashed line is seen to occur at a positive phase margin and to follow the same trend as the empirically determined phase margin adjustment rule of Eq. (13). This suggested that, provided continuous spectra shaped as in Fig. 4 looks to an operator like the augmented discrete rectangular spectra of STI, for a fixed τ_e, the crossover frequency, ω_c, is adjusted to maintain a minimum mean-squared error with changes in the input bandwidth.

In order to approach a more nearly rectangular continuous input spectrum, the input filter was modified to

$$F(s) = \frac{K}{(s+\omega_b)^4} \tag{15}$$

The results of this modification are illustrated in Fig. 5. The minimum mean squared error locus for this case has shifted to the stability limit and the bandwidth trajectories are seen to be trending toward the trajectories of the theoretical rectangular input spectrum of Fig. 3. This result indicates that a continuous white noise process may be shaped to "look-like" a rectangular spectrum with a linear filter of fourth-order or higher. This approach was utilized in the UTIAS studies of Reid and Gordon-Smith.

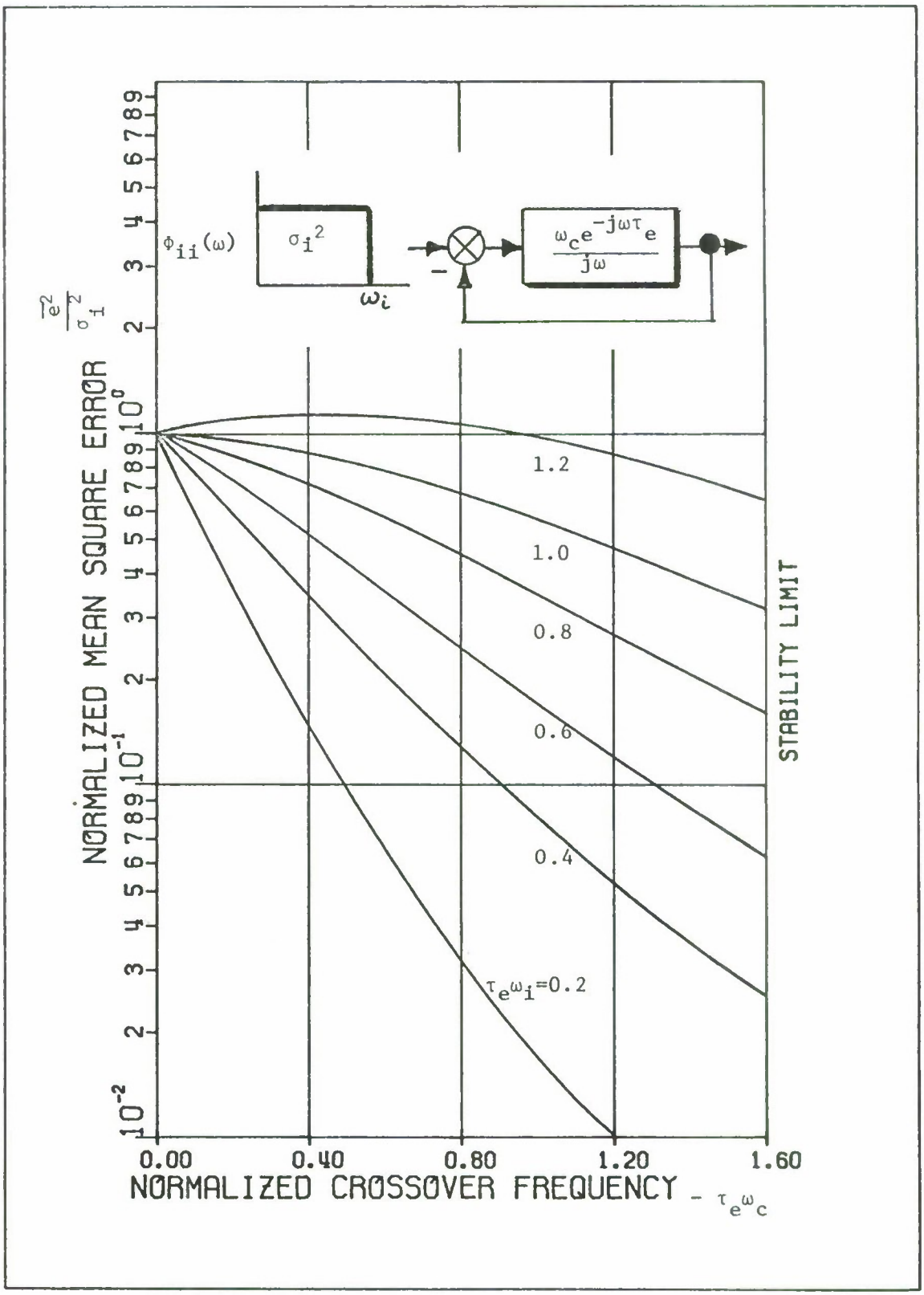


Fig. 3. Crossover Model with a Rectangular Input Spectrum

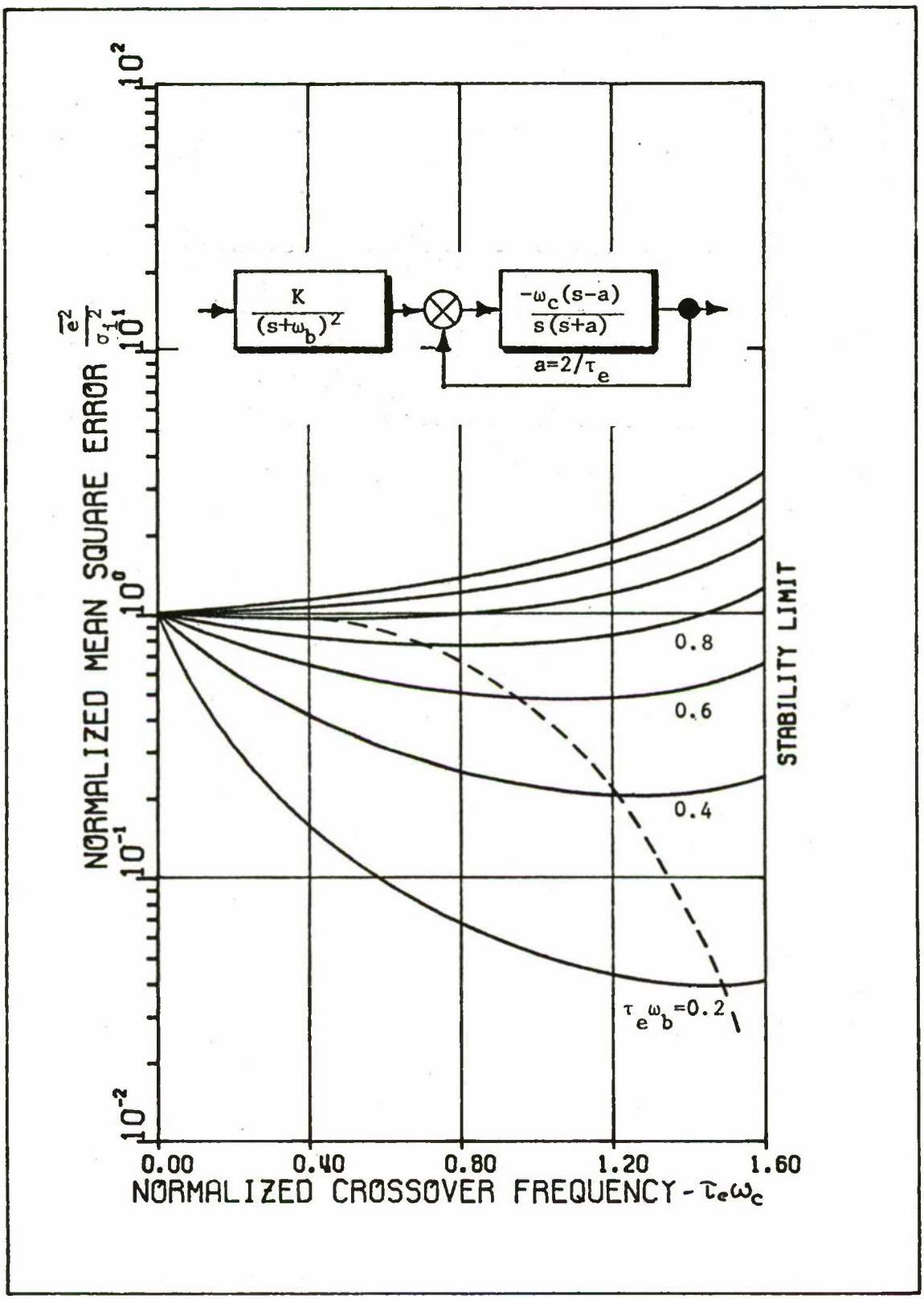


Fig. 4. Crossover Model with Second-Order Lag Filter

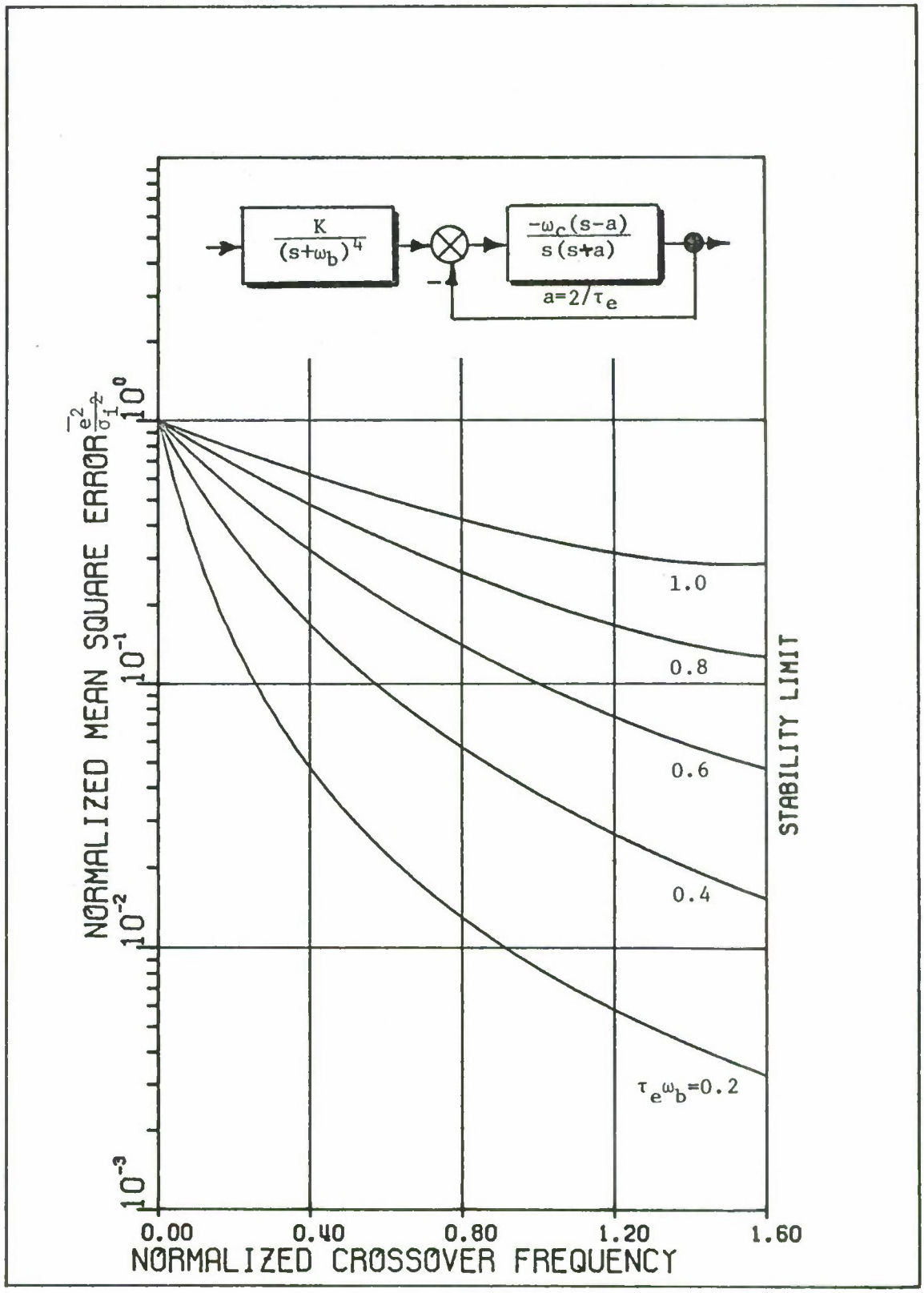


Fig. 5 Model with Fourth-Order Lag Filter

III. Experimental Results

The UTIAS Studies

Reid (Ref 13) and Gordon-Smith (Ref 4) compiled a detailed body of experimental data concerning human pilot describing functions for compensatory tracking of continuous input spectra. They worked together to devise a linear input filter which would provide continuous input spectra to match the nominal shape of the envelope of the discrete STI spectra. The input filter decided upon was of the form

$$F(s) = B(s) + C(s) \tag{16}$$

where $$B(s) = \frac{\omega_p^4}{(s^2+2\xi_p\omega_p s+\omega_p^2)^2} \frac{1}{(sT+1)}$$

and $$C(s) = \frac{R\omega_s^4}{(s^2+2\xi_s\omega_s s+\omega_s^2)^2}$$

The parameters of B(s) were chosen to sharply define the cutoff frequency of the input spectrum. The rolloff due to B(s) above the cutoff frequency occurs at a rate of 100 db/decade. The parameters of C(s) were chosen to provide the attenuated high frequency portion of the spectrum analogous to the high frequency shelf of STI. Although Gordon-Smith chose slightly different values for the parameters of F(s) from those of Reid, the resultant spectral shapes were quite similar. The UTIAS parameters are listed in Table III.

Table III
Parameters of F(s)

T	ω_s	ω_p	ξ_s	ξ_p	R
Reid (Ref 13)					
1.5	12	2	0.56	0.6	0.067
0	12	4	0.75	1.0	0.13
0	11	6	0.36	1.0	0.032
Gordon-Smith (Ref 4)					
1.33	15	1.5	0.75	0.75	-0.1
0.67	15	2.5	0.75	0.75	-0.1
0.5	15	4.0	0.75	0.70	-0.1
0.605	--	3.0	--	0.70	0

Human operator performance was investigated for three degrees of task difficulty by choosing input bandwidth cutoff frequencies of approximately 1.5, 2.5, and 4.0 rad/sec. In a fourth procedure, Gordon-Smith set the parameter R of F(s) to zero, for the mid frequency case, to investigate the effect on the removal of the high frequency shelf.

In these studies a bandwidth equivalent to that of a rectangular spectrum was defined by

$$\omega_{ie} = \frac{[\int_{-\infty}^{\infty} \Phi_{ii}(\omega)\, d\omega]^2}{\int_{-\infty}^{\infty} [\Phi_{ii}(\omega)]^2\, d\omega} \qquad (17)$$

This definition was necessary as the break frequency of an input filter cannot directly represent the bandwidth of the input power spectrum unless the filter is considered to have an infinite rolloff. The values of ω_{ie} for each forcing function spectrum are included in Table IV with the experimentally determined performance measures for controlling $K_c/(j\omega)$.

Table IV
UTIAS Performance Measures for Rate Control

ω_p rad/sec	ω_{ie} rad/sec	ω_c rad/sec	τ'_e Model sec	τ_e Actual sec	ϕ_m rad
Reid (Ref 13)					
2	1.41	3.0	0.382	0.32	0.785
4	2.8	3.9	0.196	0.24	0.803
6	4.3	3.3	0.106	0.12	1.22
Gordon-Smith (Ref 4)					
1.5	1.56	2.7	0.193	0.31	1.05
2.5	2.47	3.5	0.163	0.22	1.0
4.0	3.77	3.9	0.140	0.18	1.03
3.0*	2.26	5.7	0.192	0.19	0.472

* Without high frequency shelf.

Model Results

Fig. 6 through Fig. 12 constitute a sampling of the model results of Ref. 6 incorporating the specific linear input filter parameters used by Gordon-Smith. Figs. 6, 7, and 8 are the model representations with the experimentally determined operating regions found by Gordon-Smith for a pilot controlling a $K_c/(j\omega)$ plant in the presence of an increasing noise bandwidth. The high frequency shelf is also present in the input noise spectra for these three figures. Fig. 9 through Fig. 11 represent the theoretical model trajectories in the presence of increasing noise bandwidths for the parameters used by Gordon-Smith and with the C(s) term of Eq. (16) set to zero to eliminate the high frequency shelf. Fig. 12 has included the experimentally determined operating region for controlling a $K_c/(j\omega)$ plant with the high frequency shelf removed from the mid-frequency input noise spectrum.

The Model Trajectories

The model trajectories generated for the input spectra with a high frequency shelf were significantly different from those generated with the shelf removed. Figs. 6, 7, and 8 represent the model with the Gordon-Smith trajectories having a high frequency shelf. The trajectories of constant τ_e exhibited in these three figures exhibit definite points of double inflection, represented by a "□", within the range of interest. The inflection points for any one of these three curves represent a minimum mean-squared error operating locus. An experimentally determined operating region along this locus then would represent, in the case of controlling a $K_c/(j\omega)$ plant, an optimum pilot gain constant, K_p, in that the gain crossover frequency would be adjusted to maintain a minimum mean-squared tracking error.

The model trajectories of Figs. 9, 10, and 11 represent the Gordon-Smith parameters with the shelf removed. The model trajectories of Fig. 12 correspond in approximate input spectrum bandwidth to the trajectories of Fig. 7. The trajectories of Fig. 12 represent the Gordon-Smith experimental parameters for a mid-frequency input spectrum bandwidth without a high frequency shelf. In these figures the model trajectories of constant time delay are monotonically decreasing functions of the gain crossover frequency provided the ratio ω_{ie}/ω_c is less than one. These four model trajectories with the high frequency shelf removed are seen to be trending toward the theoretical trajectories for a rectangular input spectrum as displayed in Fig. 3.

The Experimental Results

The experimental results of Gordon-Smith are superimposed on the analytical models of Figs. 6 through 8. Additionally, Fig. 12 depicts the results of the Gordon-Smith parameters for the mid-frequency input bandwidth case with the high frequency shelf removed.

The experimentally determined operating region in the low frequency input bandwidth case of Fig. 6 can be seen to range along the locus of minimum mean-squared error. This result indicates that, in the presence of an augmented input spectral shape where the effective rectangular input bandwidth to gain crossover frequency ratio is in the region $\omega_{ie}/\omega_c \triangleq 0.5$, for controlling a fixed $K_c/(j\omega)$ plant, the operator's gain constant, K_p, is adjusted to an optimum level in the mean-squared error sense.

The results of Figs. 7 and 8 represent the mid- and high-frequency input spectrum bandwidth cases for the Gordon-Smith parameters. These results indicate that the pilot gain constant is adjusted down from optimum in the mean-squared error sense as the ratio ω_{ie}/ω_c approaches unity.

Effect of the High Frequency Shelf

The effect on human operator performance due to the removal of the high frequency shelf was reported by Gordon-Smith. Removal of the shelf caused a large reduction in score. Comparison of describing functions showed that the amplitude ratio increased with no change in phase. This represented a simple increase in pilot gain, K_p.

Figure 7 represents the analytical model with the high frequency shelf intact and Fig. 12 represents the model with the shelf removed. As mentioned previously, removal of the high frequency shelf has the effect of removing the inflections in the trajectories of constant time delay. In Fig. 7 the human operator is maintaining a gain constant which is slightly less than optimum in the mean-squared error sense. When the high frequency shelf is removed this gain constant is no longer near optimum and it can be seen in Fig. 12 that the pilot gain is increased along the constant time delay trajectory to a much lower phase margin point. The pilot is seen to switch modes from that of optimizing K_p to that of maintaining a minimum phase margin.

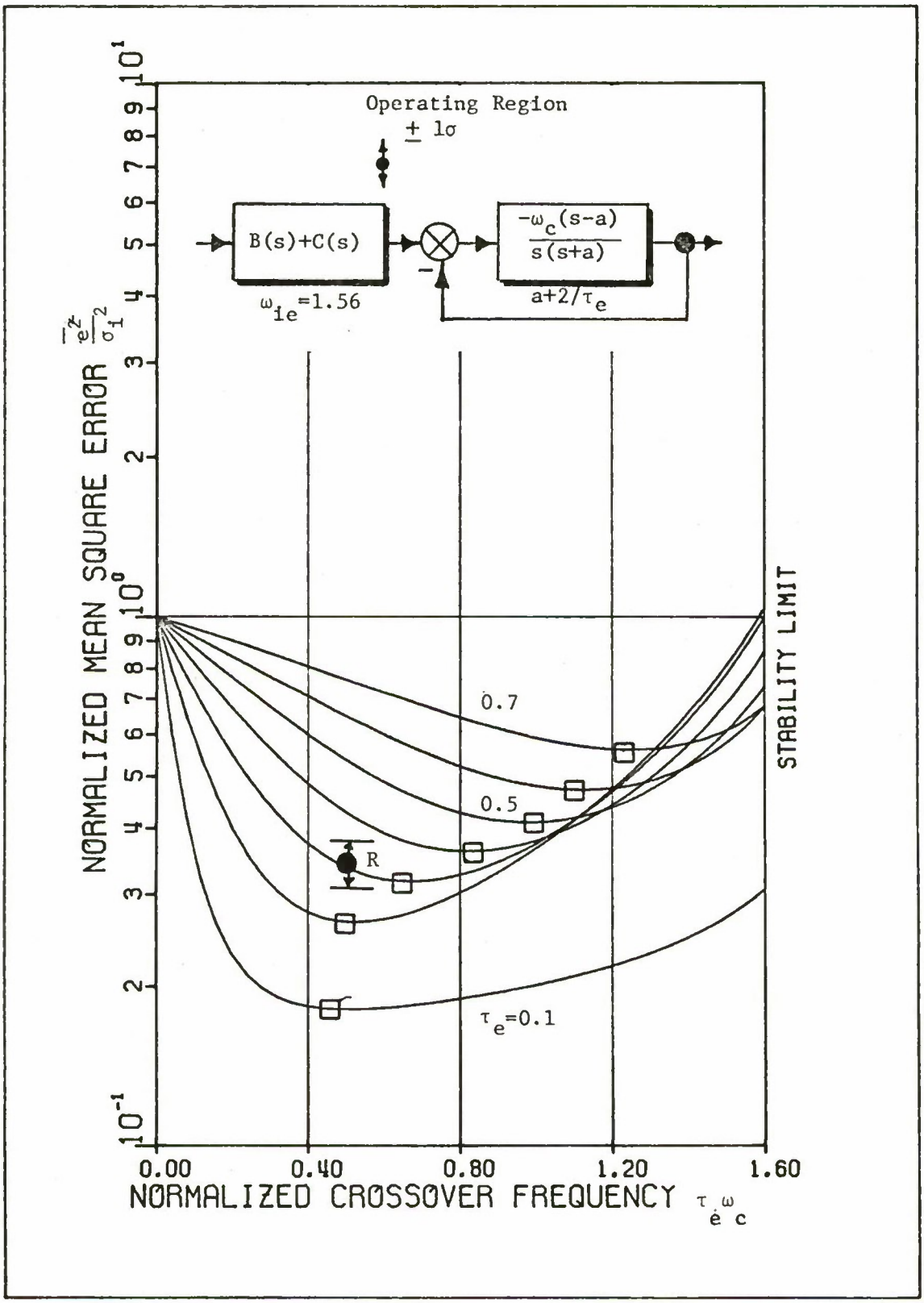


Fig. 6. Model with Gordon-Smith Parameters for ω_{ie}=1.56 Rad/Sec and with the High Frequency Shelf

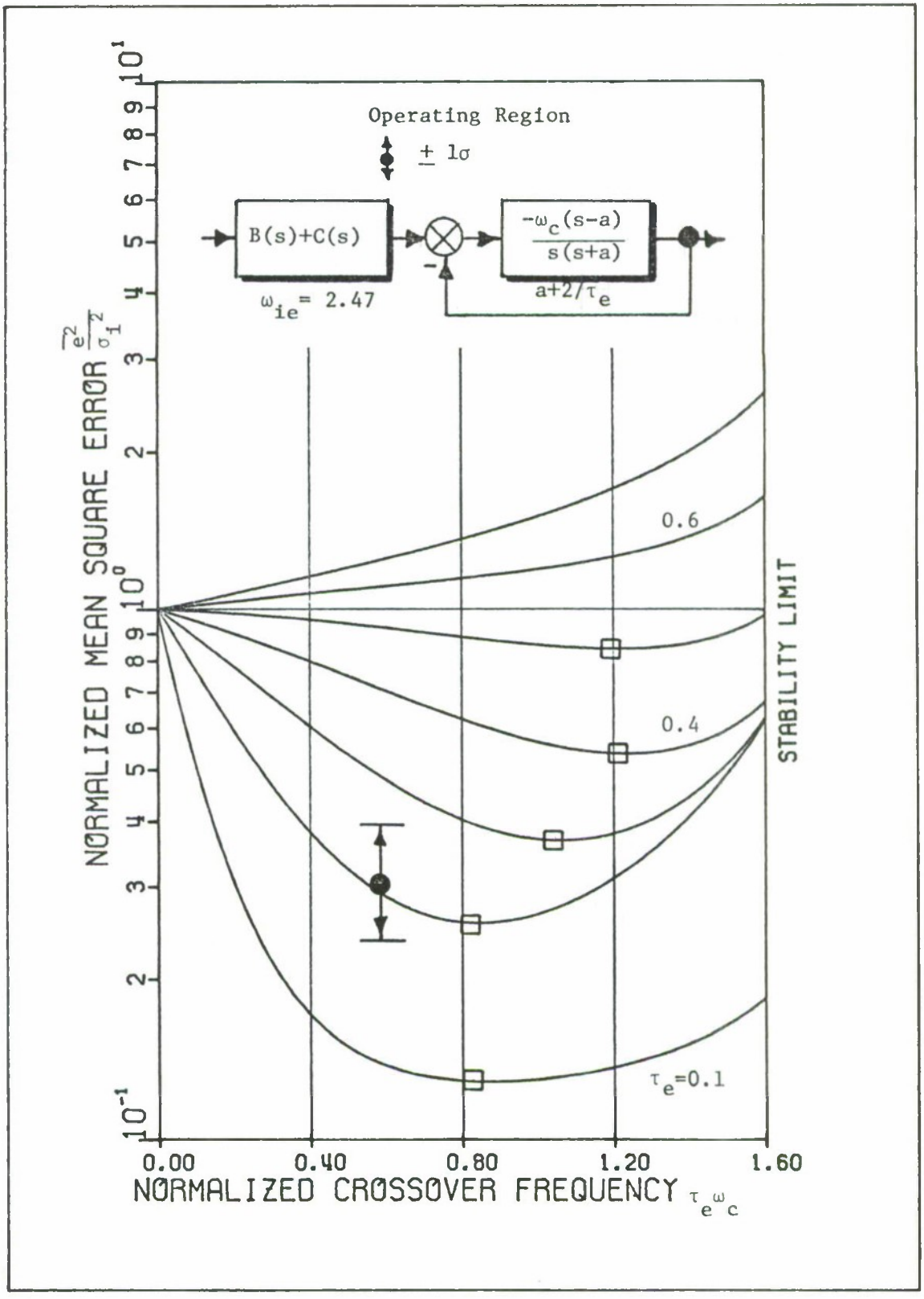


Fig. 7. Model with Gordon-Smith Parameters for ω_{ie}=2.47 Rad/Sec and with the High Frequency Shelf

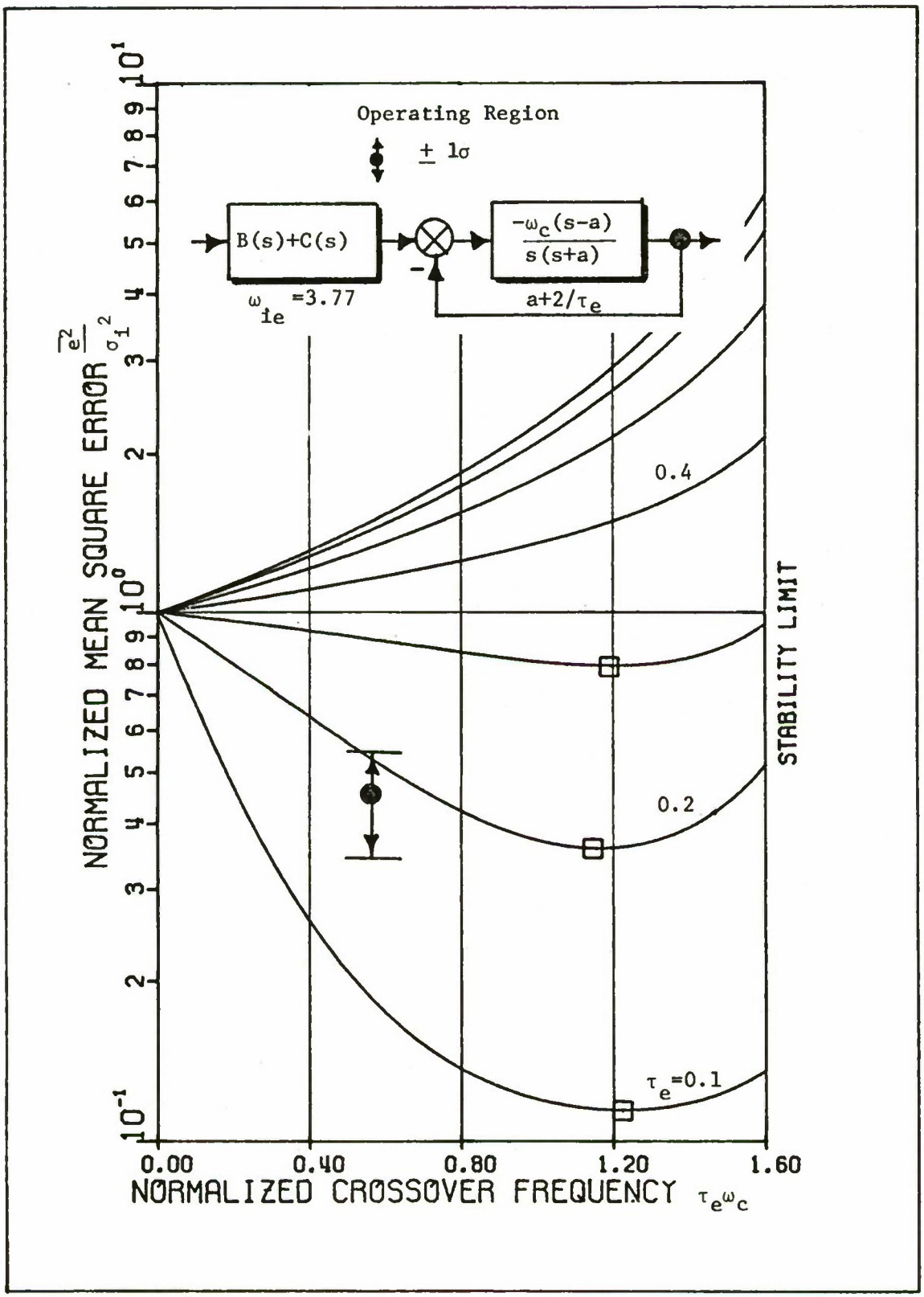


Fig. 8. Model with Gordon-Smith Parameters for ω_{ie}=3.77 Rad/Sec and with the High Frequency Shelf

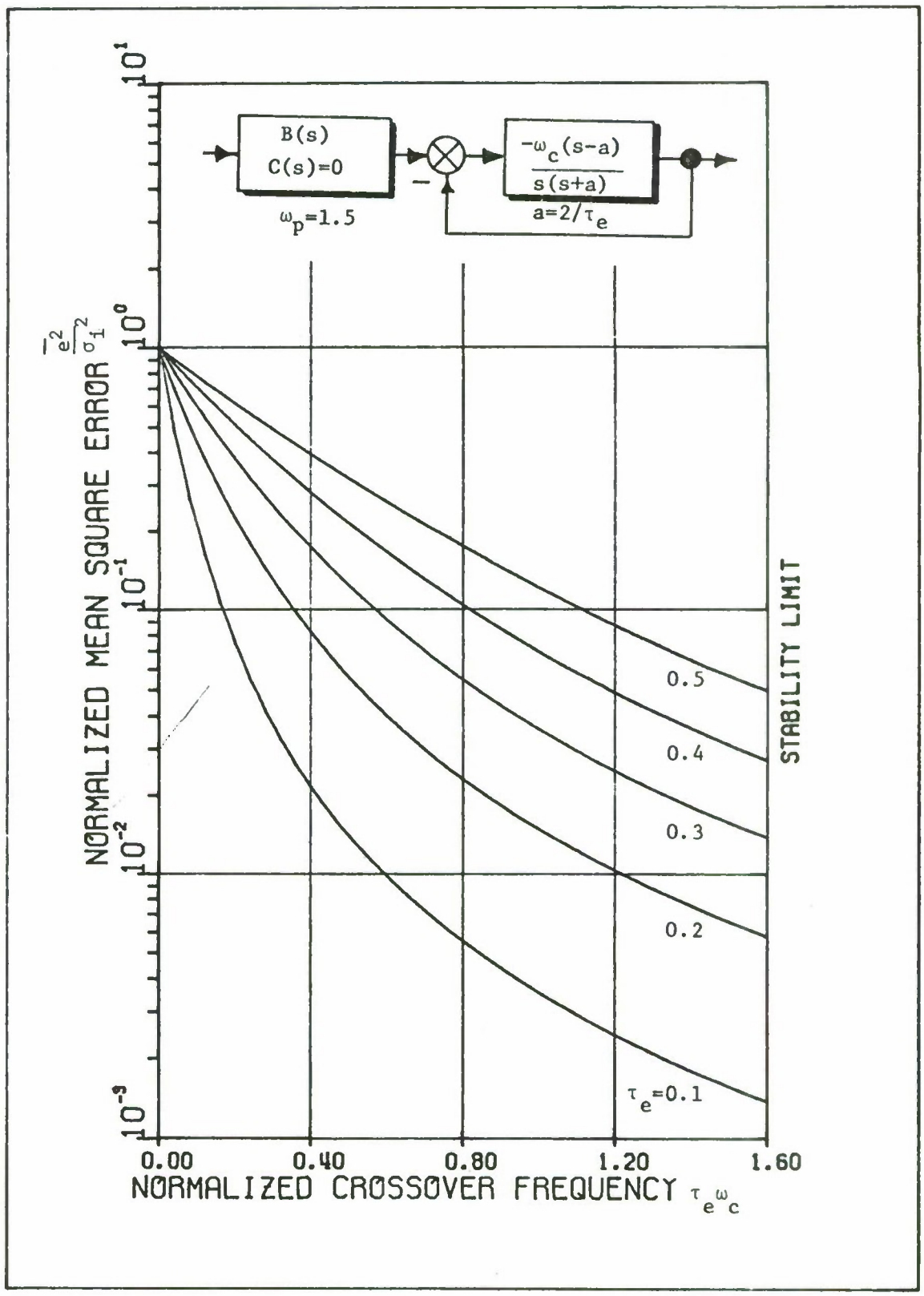


Fig. 9. Model with Gordon-Smith Parameters for ω_p=1.5 Rad/Sec and without the High Frequency Shelf

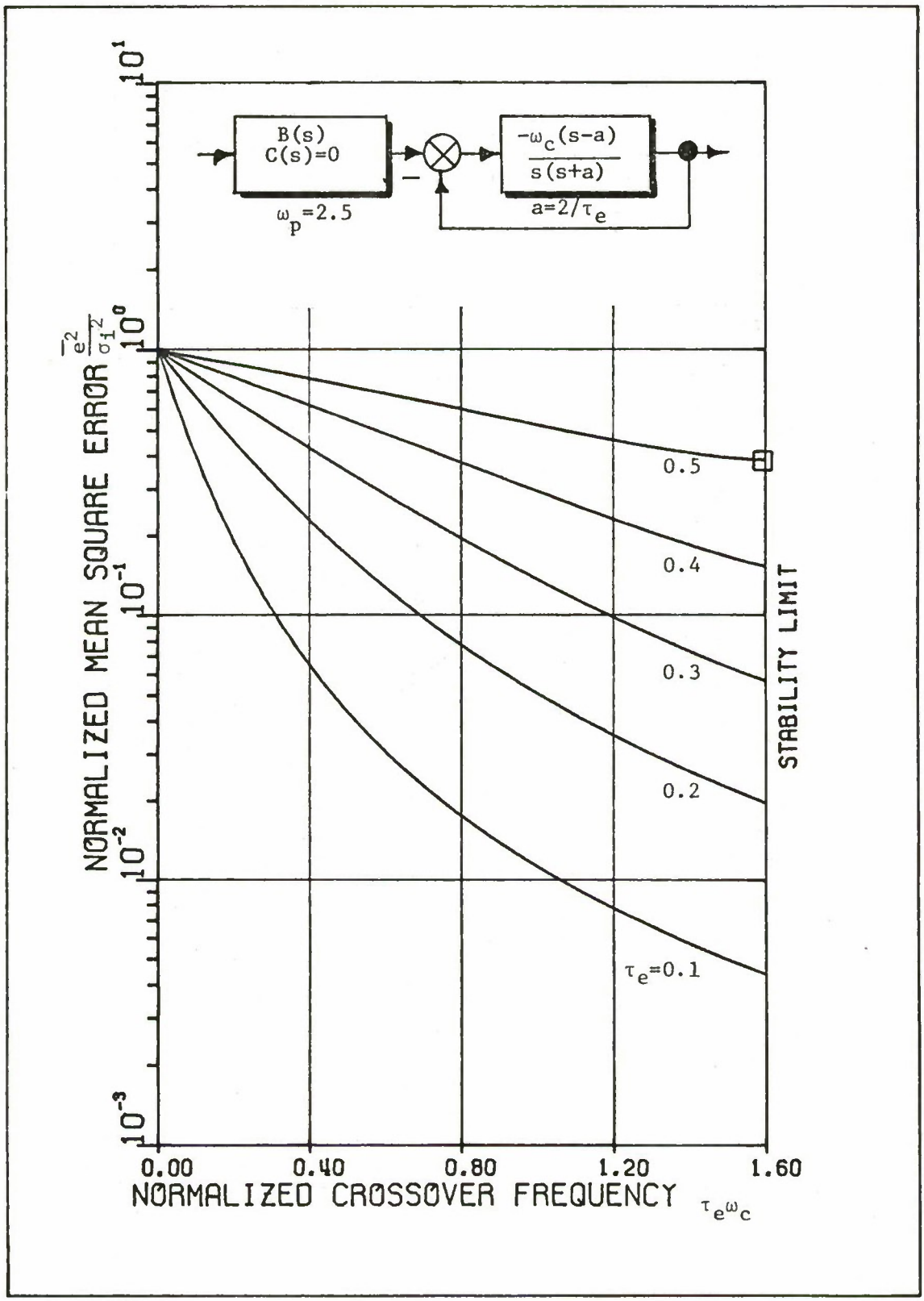


Fig. 10. Model with Gordon-Smith Parameters for ω_p=2.5 Rad/Sec and without the High Frequency Shelf

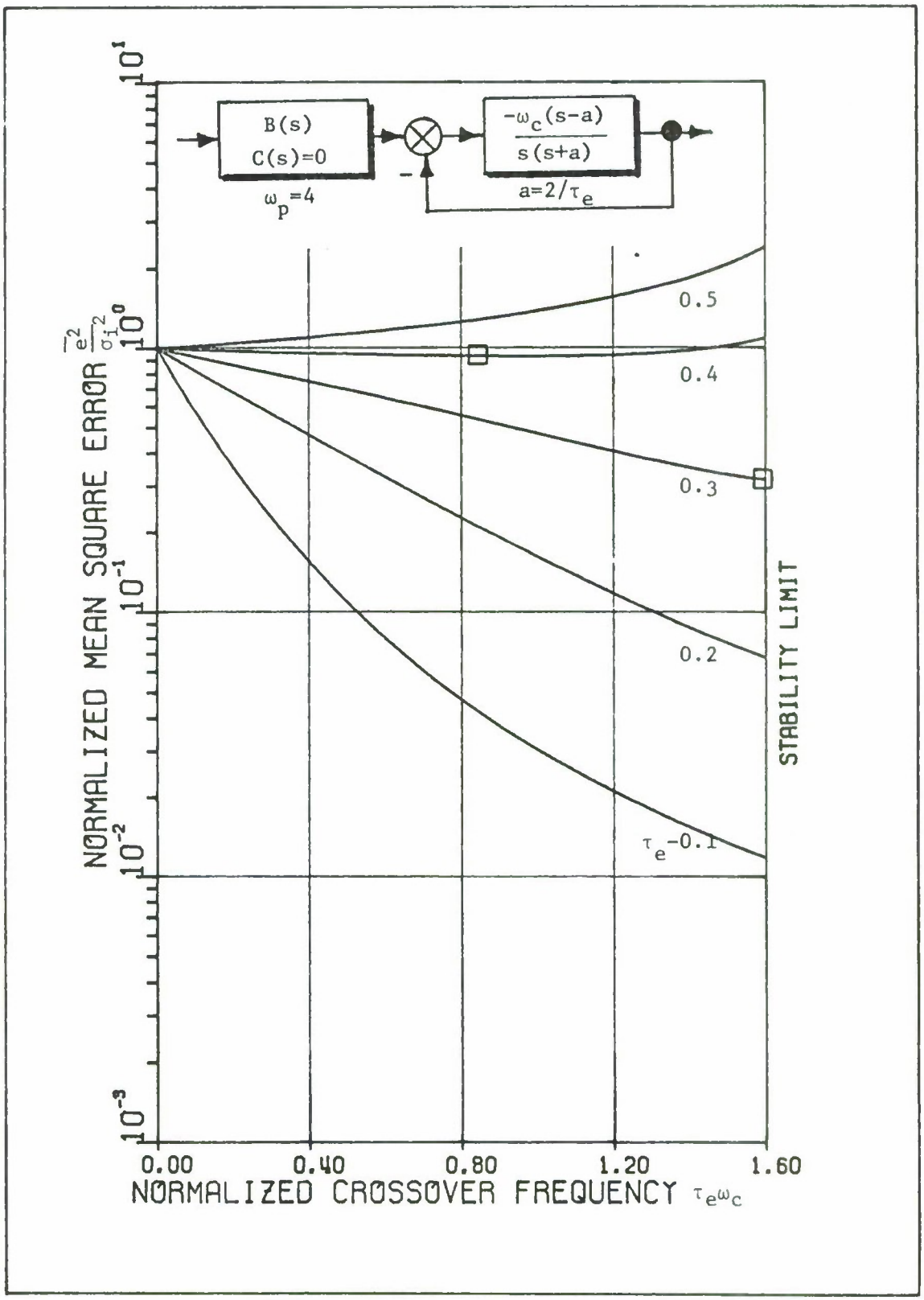


Fig. 11. Model with Gordon-Smith Parameters for ω_p=4 Rad/Sec and without High Frequency Shelf

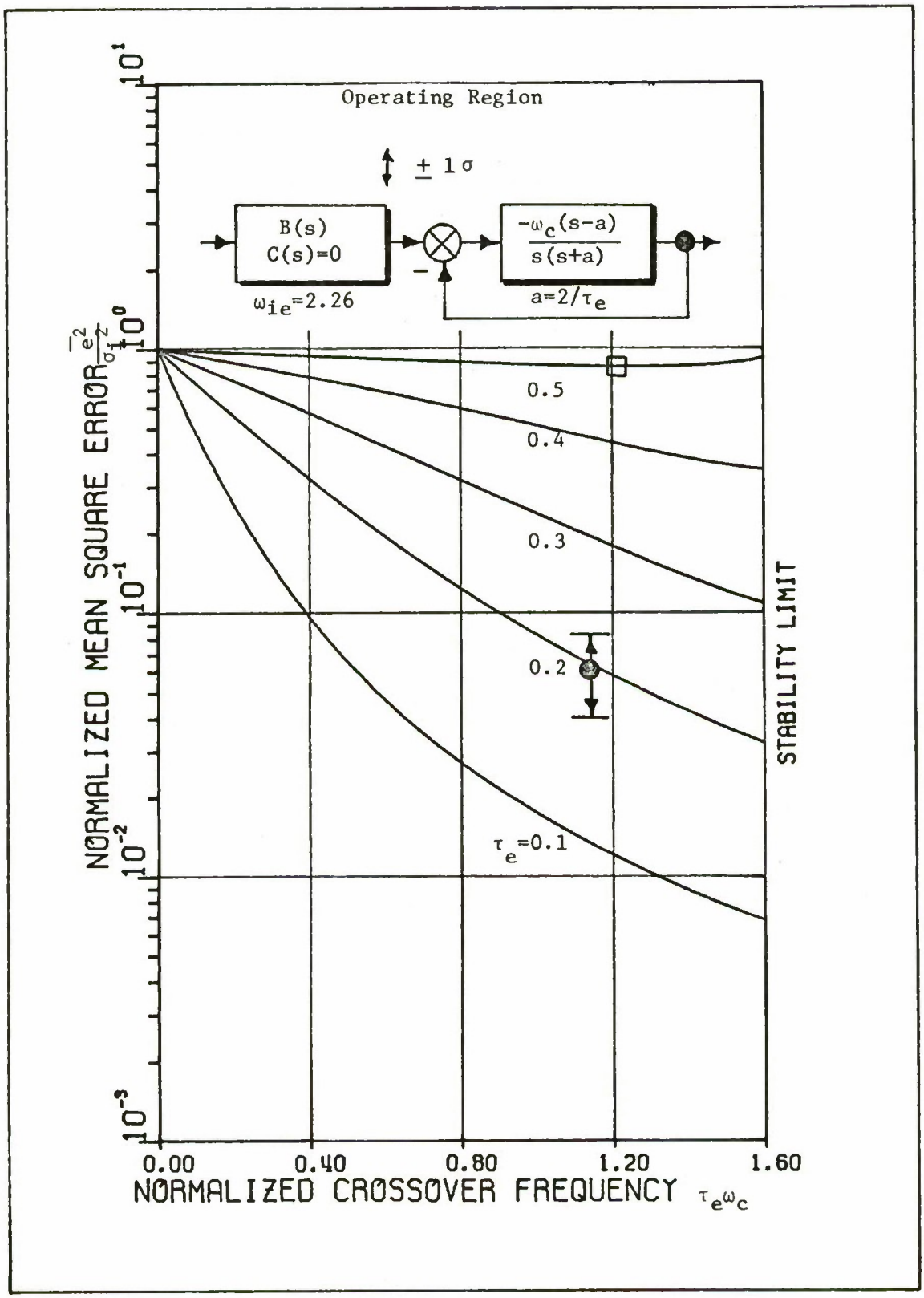


Fig. 12. Model with Gordon-Smith Parameters for ω_{ie}=2.26 Rad/Sec and without High Frequency Shelf

IV. Conclusions

The analytical model based upon both the input forcing function parameters and a polynomial representation of the crossover model, in conjunction with the concept of optimal human operator adjustments, were found to be of value in investigating the effects of input forcing function characteristics on human operator performance in a compensatory tracking task.

The model established the existence of minimum mean-squared tracking error loci, for controlling fixed element plants in the presence of continuous input spectra, which were consistent with experimental results.

Matching the experimental results of Reid (Ref 13) and Gordon-Smith (Ref 4) to the analytical model verified a predictive capability of the model. The model predicted an optimal pilot gain constant, K_p, for controlling the fixed plant of $K_c/(j\omega)$ when there was measurable energy available in the gain crossover frequency region.

The model predicted a minimum phase margin operation upon removal of energy in the gain crossover frequency region. Experimental results verified these predictions.

Based on the data ana analysis in this study, the parameter adjustment rules for continuous input spectra are different in several respects from those developed for discrete input spectra (Ref 9:19;186). By way of comparison the circa 1965 parameter adjustment rules for discrete input spectra are listed below followed by the parameter adjustment rules as determined in this study for continuous input spectra:

a. Adjustment Rule 4a: Discrete--"Closed-loop low frequency performance in operating on the forcing function is optimum in some sense analogous to that of minimizing mean-squared tracking error." Continuous--"With an augmented (shelf-type) continuous input spectra, closed-loop low frequency performance is optimum in the minimum mean-squared error sense providing the ratio ω_{ie}/ω_c is less than one."

b. Adjustment Rule 4b: Discrete--"System phase margin, ϕ_M, is directly proportional to ω_i, the forcing function bandwidth, for values of ω_i less than about 2.0 rad/sec." Continuous--"With an augmented (shelf-type) continuous input spectra, system phase margin, ϕ_M, is essentially constant and independent of the effective rectangular input bandwidth, ω_{ie}, when the ratio ω_{ie}/ω_c is less than one."

c. A new adjustment rule, 4d, should be added for continuous input spectra: "When the input spectrum approaches the theoretical rectangular spectrum equivalent to forcing white noise through a fourth-order or higher lag filter, then the system phase margin, ϕ_M, will be at a minimum consistent with good closed-loop low frequency response."

Bibliography

1. Anderson, Ronald O., Unpublished notes on minimization adjustments for a random input pilot model, Wright-Patterson Air Force Base, Ohio: Air Force Flight Dynamics Laboratory, 4 April 1969

2. Bowser, David K., On the Evaluation of Mean Square Integrals by Use of the Digital Computer. FDCC TM 65-17. Wright-Patterson Air Force Base, Ohio: Air Force Flight Dynamic Laboratory, October 1968.

3. D'Azzo, John J. and Constantine H. Houpis. Feedback Control System Analysis and Synthesis. New York: McGraw-Hill Book Company, Inc., 1966.

4. Gordon-Smith, Michael. Letter concerning An Investigation into Some Aspects of the Human Operator Describing Function while Controlling a Single Degree of Freedom. Unpublished thesis. Toronto: University of Toronto Institute of Aerospace Studies, 7 October 1969.

5. Graham, Dunstan and Duane McRuer. Analysis of Non-Linear Control Systems. New York: John Wiley and Sons, Inc., 1961.

6. Heifferon, John C. The Effects of Input Power Spectra on Human Operator Compensatory Tracking. Master's Thesis, Wright-Patterson Air Force Base, Ohio: Department of Electrical Engineering, Air Force Institute of Technology, March 1970.

7. James, H. M., N. B. Nichols, and R. S. Phillips. Theory of Servomechanisms. MIT Radiation Laboratory Series, Vol. 25, New York: McGraw-Hill Book Company, Inc., 1947.

8. Lanning, J. H., and R. H. Battin. Random Processes in Automatic Control. New York: McGraw-Hill Book Company, Inc., 1956.

9. Magdaleno, Raymond and Julian Wolkovitch. Performance Criteria for Linear Constant-Coefficient Systems with Random Inputs. ASD-TDR-62-470. Wright-Patterson Air Force Base, Ohio: Aeronautical Systems Division, Air Force Systems Command, January 1963.

10. McRuer, Duane T., et al. Human Pilot Dynamics in Compensatory Systems. AFFDL-TR-65-15. Wright-Patterson Air Force Base, Ohio: Research and Technology Division, Air Force Systems Command, July 1965.

11. McRuer, Duane T., et al. New Approaches to Human-Pilot/Vehicle Dynamic Analysis. AFFDL-TR-67-150. Wright-Patterson Air Force Base, Ohio: Air Force Flight Dynamics Laboratory, Air Force Systems Command, February 1968.

12. Papoulis, Athanasios. Probability, Random Variables, and Stochastic Processes. New York: McGraw-Hill Book Company, Inc., 1965.

13. Reid, L. D. The Measurement of Human Pilot Dynamics in a Pursuit-Plus-Disturbance Tracking Task. UTIAS Report No. 138. Toronto: University of Toronto Institute of Aerospace Studies, April 1969.

14. Teasdale, R. D. "Time Domain Approximation by Use of Padé Approximants." IRE Convention Record, 1953, Part 5--Circuit Theory, pp. 89-94.

15. Truxal, R. D. Automatic Feedback Control Synthesis. New York: McGraw-Hill Book Company, Inc., 1955.

ON THE VARIANCE OF THE BICYCLE RIDER'S BEHAVIOR *

A. van Lunteren H.G. Stassen

Man-Machine Systems Group of the Laboratory for Measurement and Control, Department of Mechanical Engineering, Delft University of Technology, The Netherlands.

0. Abstract

The behavior of a rider stabilizing a bicycle simulator has been studied. The simulator used demonstrates a reasonable similarity to a normal bicycle; the forward motion is missing, however, its effects on the dynamics of the simulator are taken into account.

The behavior of the cyclist has been described by the describing functions between the input of the rider, viz. the frame angle, and the outputs, viz. the rotations of handle bar and upper body. The parameters of the model in this way obtained were determined using an on-line open loop parameter estimation method. For low values of the remnants the bias due to the use of an open loop method in a closed loop system is small.

It has been found that the behavior of the rider is time-independent over at least five minutes. Furthermore, if σ_a is the mean value of the standard deviation of the parameters for one subject within one test, if σ_b is the mean value of the standard deviation for one subject over a number of tests, and if σ_c is the mean value of the standard deviation for a group of subjects, then the relation between these quantities can be approximated by $\sigma_a : \sigma_b : \sigma_c = 1 : 2 : 3$.

1. Introduction

The origins of the bicycle stabilization study lie in the field of those man-machine relations where, as distinct from experiments with astronauts or pilots, the operator comes from a relatively non-selected population, since the bicycle is one of the most popular means of transport in

* To be presented at the 6th Annual Conference on Manual Control, april 7-9, 1970, Wright-Patterson AFB, Ohio.

The Netherlands. Another important aspect is that the bicycle is an unstable system, so to stabilize it, the rider is forced to pay close attention.

The rider-bicycle system is a multiloop control system as well as a multimodality system in the sense of McRuer's definition [1]. Moreover, it has a time-varying character. In investigating only the stabilization phenomenon, the rider can be described by a model with one input (the angle between vertical and frame) and two outputs (the control actions of upper body and handle bar). Investigations of Young [2] and Donaldson [3], describing the balancing of an inverted pendulum using visual and/or motion cues are based on similar considerations. By placing the rider in normal traffic situations, the stabilization phenomenon can be used as a critical-instability task for secondary work load research as mentioned by Jex [4].

The particular development reported in this paper started with an analysis of the rider by using correlation techniques to describe his stabilization task [5]. It was shown that the rider could be described by a three-term controller with a time delay [5; 6]. Later on, an on-line calculation of the parameters of the mathematical description of the rider was used. An analysis of the course following task is under study.

A detailed description of the work reported in this paper is given in the Annual Report 1969 of the Man-Machine Systems Group [7].

2. The bicycle simulator

Based on a study of bicycle dynamics with reference to the stabilization phenomenon, published among others by Whipple in 1899 [8], a laboratory bicycle simulator was built. The following simplifications were introduced:

a. Only the stabilization task on a straight course is considered.
b. The coupling between the rotation of frame and handle bar is neglected.
c. The body is considered to be consisting of two solid parts, viz. the upper body and lower body.
d. Only small deviations about the equilibrium position are considered (linearization).
e. Only velocities in the direction of motion between 10 and 20 km/h will be simulated.

Under the assumptions mentioned above the rider-bicycle system is described by the linear differential Eq.1 derived from the moments acting on the frame about the ox axis through the contact points between wheels and

ground (as modified from [8]):

$$\underbrace{Av\dot{\zeta}_1(t)}_{\text{GYROSCOPIC MOMENT}} + \underbrace{Bv^2\zeta_1(t)}_{\text{CENTRIFUGAL MOMENT}} = \underbrace{I\ddot{\varphi}(t) - Sg\varphi(t)}_{\text{FRAME TILT}} + \underbrace{(I^*+z_1S^*)\ddot{\zeta}_2(t) - S^*g\zeta_2(t)}_{\text{UPPER BODY}} \quad (1)$$

Fig. 1 shows the axis systems used; the quantities are defined as follows:
$\zeta_1(t)$ is the angle between handle bar and frame;

$\zeta_2(t)$ is the angle between upper body and frame;

$\varphi(t)$ is the angle between the vertical oz axis and the frame; I and S are the moment of inertia and the static moment respectively of frame, wheels and lower part of the human body about the ox axis; I^* and S^* are the moment of inertia and the static moment respectively of the upper body about an o^*x^* axis parallel to the ox axis and intersecting the saddle; z_1 is the distance between the ox axis and the o^*x^* axis; A and B are constants, and v is the forward velocity.

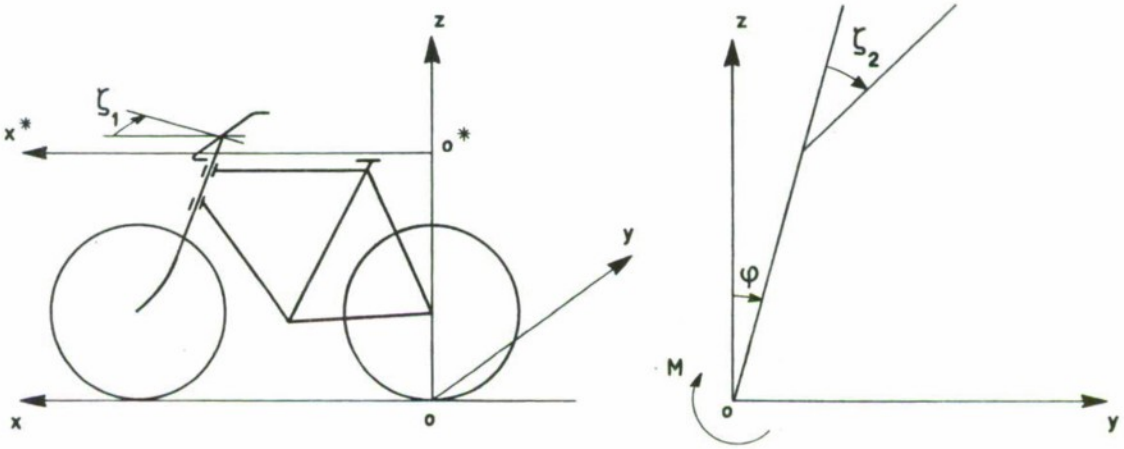


Figure 1: Definition of the variables of the bicycle simulator.

Starting from the averaged values of weights and dimensions of the various parts of the human body [9] and considering for instance a man with a weight of 63 kg and a height of 1.77 m., the numerical values for the constants of Eq. 1 are: I = 109 kgm^2; Sg = 940 Nm;

$I^* + z_1S^* = 23.9$ kgm^2; S^*g = 162 Nm; A = 25 kgm and B = 57.5 kg.

The laboratory bicycle simulator demonstrates a reasonable similarity to a normal bicycle; in general a rider is able to stabilize the simulator after a few minutes of practice. The forward motion is missing, but its effects on the simulator are taken into account, see Fig. 2.

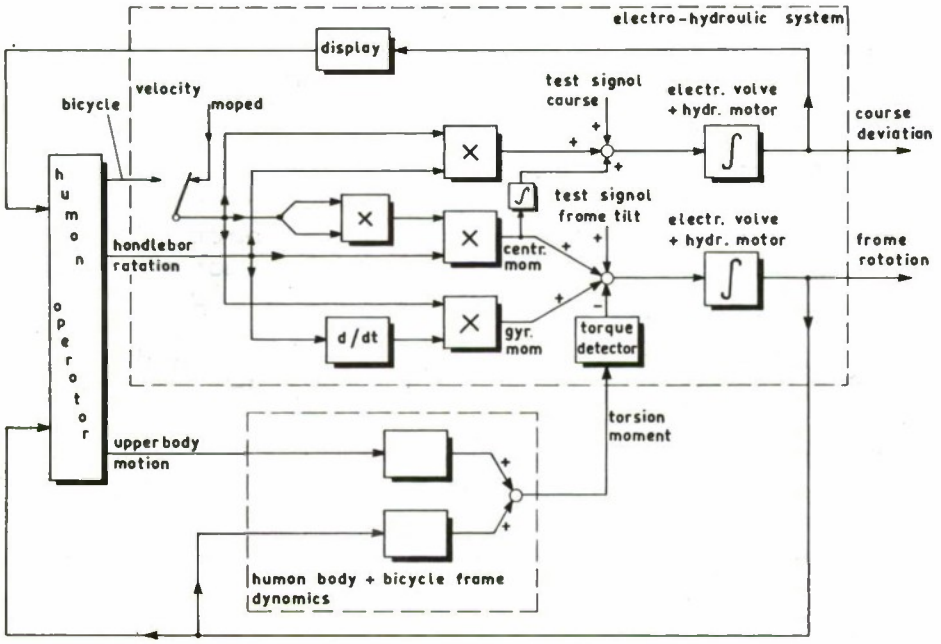


Figure 2: Blockdiagram of the bicycle simulator and its interaction with the human operator.

The rotation of the shaft ox on which the frame is mounted, is accomplished by means of an electro-hydraulic servomotor. From the difference between the electrically generated gyroscopic and centrifugal moments, and the moment about the ox axis, measured by means of a torsion transducer, the servomotor is controlled. The characteristics of the bicycle model can be changed readily by varying the electronic components only. The forward velocity can be generated by an electromotor or by the rider. In the first case the simulator is used as a motorized bicycle often called a moped; in the latter as a normal bicycle. Recently a visual display unit projecting a line representing the actual course on a screen has been added. This unit is controlled by the handle bar angle and the forward velocity, and consists of an electrically controlled hydraulic servomechanism driving a turntable with a projector. For a constant velocity v the

course deviation can be described by

$$\xi(t) = \int_0^t \{0,47\ v\psi(\tau_1) + 0,85 \int_0^{\tau_1} v^2\psi(\tau_2)d\tau_2\}d\tau_1. \qquad (2)$$

To measure human performance, unpredictable test signals simulating sidewind variations and small course deviations can be introduced into the system. The simulator is shown in Fig. 3.

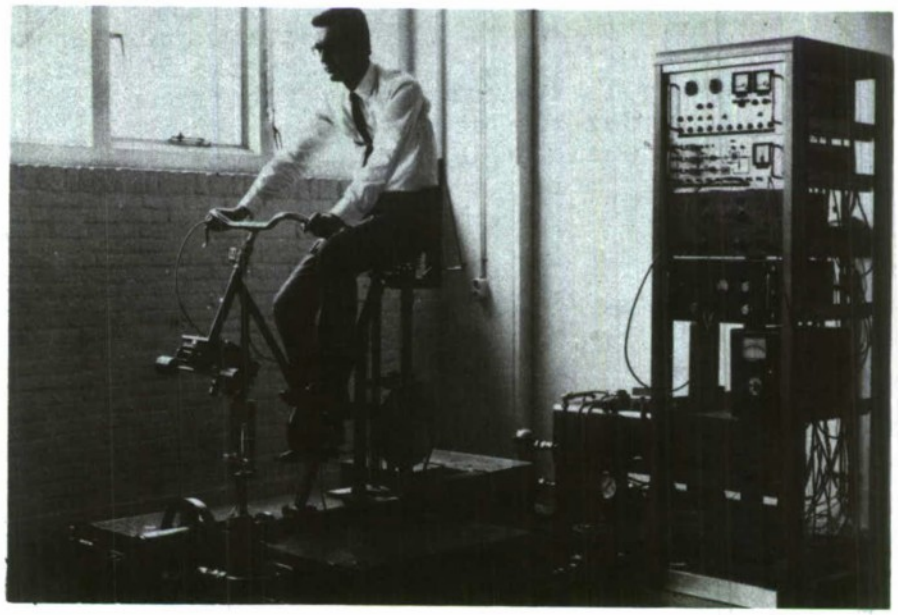


Figure 3: The bicycle simulator with the subject in a stabilization task.

3. The modeling of the rider's behavior in a stabilization task

A first approach to determine the characteristics of the rider executing a stabilization task was based on correlation methods. The necessary covariance functions were calculated by means of a special correlation method, the polarity coincidence correlation method [10]. The test signals used in the bicycle experiment have a gaussian distribution; this does not imply that the other signals such as the control variables of the rider will have the

same statistical properties, since the distributions of the operator's remnants are unknown.

The choice of the structure of the model was based on outcomes of correlation studies as well as on the following considerations. Human behavior is strongly dependent on the dynamics of the system to be controlled [11]. Investigations by Tustin [12] and Ragazzini [13] showed that the behavior of a human operator in a man-machine system could be described by a linear model consisting of a three-term controller with a delay time and, in the context of the describing function method [14], an additive remnant. Other linear models, described by a transfer function with 1 zero, 2 poles and a delay time, have been suggested [11; 15; 16; 17]. Restricted by the computer available, a three-term controller model based on a direct method for on-line parameter estimation was applied. If the angle φ is considered as the input to the subject informing him about the state of the bicycle, and if the variables ζ_i are considered as the control outputs of the subject ($i = 1$ for the motions of the handle bar; $i = 2$ for those of the upper body), then the form of the human transfer function will be:

$$H_i(s) = (\alpha_{i1} + \alpha_{i2}/s + \alpha_{i3}s) \exp\{-\tau_i s\}, \tag{3}$$

where s is the Laplace operator. The parameters α_{i1}, α_{i2} and α_{i3} represent the proportional, integral and derivative constants respectively, and τ_i the delay time. To complete the describing function model of the bicycle rider the remnants $n_i(t)$ are introduced.

The parameter estimation method to be derived here, is in fact only valid for the identification of an unknown transfer function, for instance that of a human operator, in an open loop system. Consider a model with the transfer function of Eq. 3, and with an input x(t) then the output $\zeta_i^*(t)$ of the model will be:

$$\zeta_i^*(t) = \alpha_{i1}x(t-\tau_i)+\alpha_{i2}\frac{1}{p}\{x(t-\tau_i)\}+\alpha_{i3}p\{x(t-\tau_i)\}, \tag{4}$$

where p is the Heaviside operator d/dt. By defining the quantities $u_1(t) = x(t)$, $u_2(t) = \frac{1}{p}\{x(t)\}$ and $u_3(t) = p\{x(t)\}$ and by using a vector notation Eq. 4 can be rewritten as:

$$\zeta_i^*(t) = \underline{\alpha}_i^T \underline{u}(t-\tau_i) = \underline{u}^T(t-\tau_i)\underline{\alpha}_i. \tag{5}$$

Furthermore, the error $e_i(t)$ between the output $\zeta_i(t)$ of the human operator and the output $\zeta_i^*(t)$ of the model can be defined as:

$$e_i(t) = \zeta_i(t)-\zeta_i^*(t) = \zeta_i(t)-\underline{\alpha}_i^T \underline{u}(t-\tau_i) = \zeta_i(t)-\underline{u}^T(t-\tau_i)\underline{\alpha}_i.$$

Starting from the mean squared error criterion

$$E_i(\theta) = \int_0^\theta e^2(t)dt, \tag{7}$$

the parameters $\underline{\alpha}_i$ and τ_i can be computed by minimizing the quantity $E_i(\theta)$ with respect to the quantities $\underline{\alpha}_i$ and τ_i. Minimizing with respect to the parameters $\underline{\alpha}_i$ yields:

$$\text{grad}_{\underline{\alpha}_i}\{E_i(\theta)\} = \text{grad}_{\underline{\alpha}_i}\{\int_0^\theta e_i^2(t)dt\} =$$

$$= 2\int_0^\theta e_i(t)\ \text{grad}_{\underline{\alpha}_i}\{e_i(t)\}dt = \underline{0}. \tag{8}$$

The set of so-called sensitivity functions $\text{grad}_{\underline{\alpha}_i}\{e_i(t)\}$ can be written as:

$$\text{grad}_{\underline{\alpha}_i}\{e_i(t)\} = \text{grad}_{\underline{\alpha}_i}\{\zeta_i(t) - \underline{u}^T(t-\tau_i)\underline{\alpha}_i\} = -\underline{u}^T(t-\tau_i). \tag{9}$$

By inserting Eqs 6 and 9 in Eq. 8 the following relation is found:

$$\int_0^\theta \zeta_i(t)\underline{u}^T(t-\tau_i)dt = \underline{\alpha}_i^T \int_0^\theta \underline{u}(t-\tau_i)\underline{u}^T(t-\tau_i)dt. \tag{10}$$

For $\tau_i \ll \theta$ the following approximation is allowed:

$$\int_0^\theta u_j(t-\tau_i)u_k(t-\tau_i)dt \approx \int_0^\theta u_j(t)u_k(t)dt. \tag{11}$$

Furthermore, by defining the vector $\underline{b}$ and the matrix A in such a way that:

$$b_{ik} = \int_0^\theta \zeta_i(t)u_k(t-\tau_i)dt; \tag{12}$$

$$a_{jk} = a_{kj} = \int_0^\theta u_j(t)u_k(t)dt, \tag{13}$$

Eq. 10 can be rewritten as:

$$\underline{b}_i^T = \underline{\alpha}_i^T A = \underline{\alpha}_i^T A^T \quad \text{or} \quad \underline{b}_i = A\underline{\alpha}_i, \tag{14}$$

$$\text{so that } \underline{\alpha}_i = A^{-1}\underline{b}_i. \tag{15}$$

The quantity $E_i(\theta)$ can be expressed as:

$$E_i(\theta) = \int_0^\theta e_i^2(t)dt = \int_0^\theta e_i(t)\ \zeta_i(t)dt - \{\int_0^\theta e_i(t)\underline{u}^T(t-\tau_i)dt\}\underline{\alpha}_i. \tag{16}$$

From Eqs 8 and 9 then it follows that:

$$E_i'(\theta) = \int_0^\theta e_i(t)\ \zeta_i(t)dt = \int_0^\theta \zeta_i^2(t)dt - \underline{\alpha}_i^T \int_0^\theta \zeta_i(t)\underline{u}(t-\tau_i)dt \tag{17}$$

where $E_i'(\theta) = E_i(\theta)$ minimized with respect to $\underline{\alpha}_i$, or

$$E_i'(\theta) = \int_0^\theta \zeta_i^2(t)dt - \underline{\alpha}_i^T\underline{b}_i = \int_0^\theta \zeta_i^2(t)dt - \underline{b}_i^T A^{-1}\underline{b}_i. \tag{18}$$

Minimization of $E_i(\theta)$ with respect to τ_i is less easy than with respect to $\underline{\alpha}_i$. Therefore, the following strategy is applied. A number of values is chosen for the quantity τ covering the range in which the delay times τ_i are expected. For each of the values of τ the quantities:

$$\underline{\alpha}_i'(\tau) = A^{-1}\underline{b}_i(\tau);$$

$$E_i'(\theta,\tau) = \int_0^\theta \zeta_i^2(t)dt - \underline{b}_i^T(\tau)A^{-1}\underline{b}_i(\tau)$$

are calculated. Furthermore, the functions $\underline{\alpha}_i'(\tau)$ and $E_i'(\theta,\tau)$ are approximated by a set of polynomials of the quantity τ. Consecutively, the value τ_i for which $E_i'(\tau)$ is minimal, is determined and inserted in the polynomials for $\underline{\alpha}_i'(\tau)$, thus yielding all desired parameters. The elements of the vectors $\underline{b}_i'(\tau)$ for a number of previously chosen values of τ and the elements of A (which do not depend on τ) can be computed on-line. Note that from Eq. 8 it follows that the error e(t) is orthogonal with respect to the sensitivity functions $\underline{u}(t-\tau_i)$ over the chosen interval $0 < t < \theta$. If a proper structure is chosen for the model it also will be uncorrelated with the input x(t) so that e(t) will be equal to the remnant $n_i(t)$. The following three comments should be made:

a. The uncertainty of the parameter estimation due to the finite time of observation decreases with a lengthening in time. However, the uncertainty due to the time-varying behavior then increases. The optimal observation time has to be determined experimentally.

b. The sensitivity of α_{ij} ($i = 1,2$; $j = 1,2,3$) to τ_i in the neighborhood of τ_i also has to be checked in practice, in this way getting information on the error in α_{ij} which results from an error in τ_i.
c. The method is valid to determine the transfer function in an open loop system. If it is applied in a closed loop system a bias due to the remnant is introduced.

4. Computer program and instrumentation

A PDP8 has been programmed in such a way that the quantities $\alpha_{ij}(\tau)$ and $E_i'(\tau)$ for 5 values of τ can be computed. The quantities $E_i'(\tau)$ are normalized with respect to the variances $\sigma_{\zeta_i}^2$ of the outputs $\zeta_i(t)$. The running averages of the quantities σ_{ζ_i}, σ_φ as well as the quantities $\alpha_{ij}(\tau)$ and $E_i'(\tau)$ for the 5 values of τ selected are determined every minute over the preceding 5 minutes of observation. At the time of this investigation the PDP8 available had a 4k memory, therefore it was not yet possible to implement the procedure for polynomial interpolation. So this operation was performed afterwards. To show the sensitivity of α_{ij} to τ_i, a measure for the sensitivity, $\frac{\Delta\alpha_{ij}}{\Delta\tau}$ was calculated. Figure 4 gives a block diagram of the method applied.

As a test signal for the man-bicycle system white noise was used filtered by means of a 4th order band-pass filter with a band width between 0.02 and 5 Hz. In all experiments described in this paper the simulator was used as moped having a velocity of 15 km/h.

To remove drift and noise from the computer inputs, the quantities $\varphi(t)$ and $\zeta_i(t)$ were filtered by a first-order band-pass filter between 0.02 and 5 Hz. From the angle $\varphi(t)$ the sensitivity functions $u_j(t)$ were determined by analog computation. The differentiating network had a cut-off frequency of 50 Hz, the integrating network one of 0.002 Hz. The sample frequency for each variable was 30 Hz.

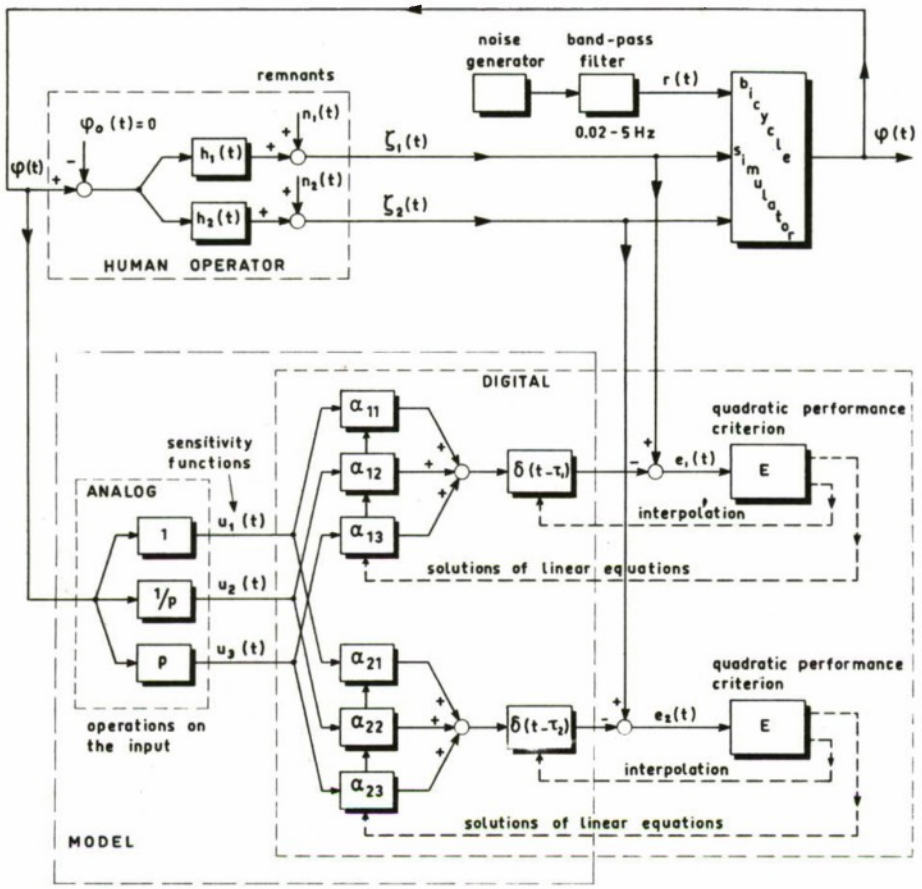


Figure 4: Block diagram of the on-line computation of the parameters of the mathematical model of the rider.

5. Practical investigations

From the cross-covariance functions between the test signal and the input and outputs of the rider respectively, the transfer functions were computed. The transfer functions are unbiased. Based on the auto-covariance function of the input as well as the cross-covariance functions between input and outputs of the rider, the transfer function can be calculated as well. However, in this case the transfer functions are biased. The same bias error will be found in using the

parameter estimation program. To show the influence of the bias, in Fig. 5 the three transfer functions are given.

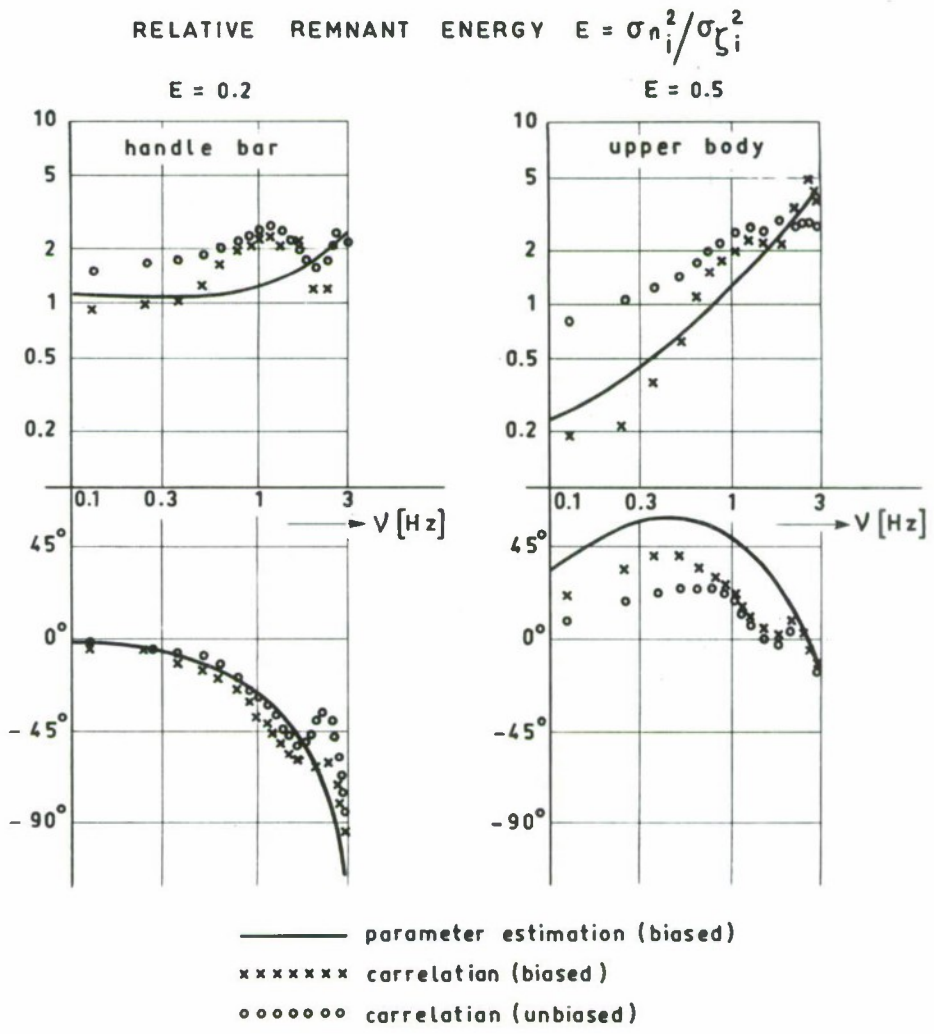


Figure 5: Transfer functions of a human operator stabilizing a bicycle simulator.

The figure shows the difference between the first method, in particular suitable for closed loop measurements, and the last two methods, in fact only valid for the open loop case. The following conclusions can be drawn:

a. The bias effect is small for low values of the relative remnant energy

$$E = \sigma_{n_i}^2/\sigma_{\zeta_i}^2, \quad E \lesssim 0.2.$$

b. For higher values, for instance E = 0.5, the bias does not change the structure of the model. Only a change in the parameters, particularly in the gain factor, occurs.
c. The outcomes obtained by the correlation techniques and the biased parameter estimation method indicate that the integrative action is missing in this particular investigation.

The further research has been executed with the parameter estimation program for the three-term controller model. The first point of interest was the variance of the behavior of the rider within one experiment. For this purpose, two male subjects A and B performed a testrun over 30 minutes; all revelant signals were recorded on magnetic tape. Afterwards the parameters were determined over 30 intervals of one minute each. From these 30 values for the 16 quantities α_{ij}; τ_i; E_i and $\frac{\Delta\alpha_{ij}}{30\Delta\tau}$ the mean value and the standard deviations σ_1 were calculated. By dividing the testruns in units of 2,3,4 and 5 minutes again the standard deviations σ_2, σ_3, σ_4, and σ_5 were determined for the quantities just-mentioned. Furthermore, the standard deviations normalized to the one minute case, σ_k/σ_1 for k = 1,2,3 4 and 5, were calculated. In order to compress the results obtained, the mean value and the standard deviation over the 16 quantities of the normalized standard deviations σ_k/σ_1 were determined for each value k. The results presented in Fig. 6 show that the behavior of the rider can be regarded as time-invariant

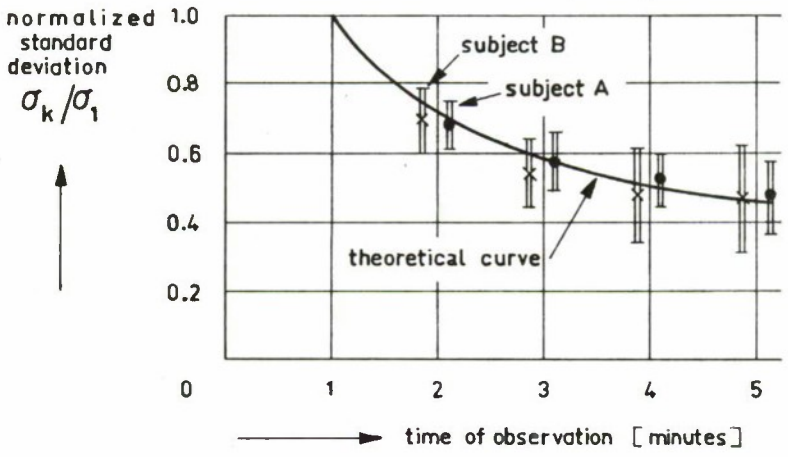


Figure 6: Influence of the observation time on the standard deviation related to the behavior of the bicycle rider.

for an observation time of at least 5 minutes. The mean values μ and the variance σ_5 of the quantities mentioned are given in Table 1. The table shows that

HANDLE BAR									
SUBJECT		α_{11}	α_{12}	α_{13}	E_1	τ_1	$\frac{\Delta\alpha 11}{30\Delta\tau}$	$\frac{\Delta\alpha 12}{30\Delta\tau}$	$\frac{\Delta\alpha 13}{30\Delta\tau}$
A	μ	0.95	-0.014	0.14	0.27	0.160	-0.04	0.001	0.029
	σ_s	0.05	0.009	0.03	0.06	0.011	0.01	0.001	0.002
B	μ	1.15	0.007	0.12	0.19	0.157	-0.027	0.001	0.034
	σ_s	0.03	0.005	0.02	0.01	0.008	0.003	0.0005	0.004

UPPER BODY									
SUBJECT		α_{21}	α_{22}	α_{23}	E_2	τ_2	$\frac{\Delta\alpha 21}{30\Delta\tau}$	$\frac{\Delta\alpha 22}{30\Delta\tau}$	$\frac{\Delta\alpha 23}{30\Delta\tau}$
A	μ	-0.42	-0.01	-0.13	0.46	0.087	0.038	-0.002	-0.015
	σ_s	0.06	0.01	0.01	0.07	0.009	0.002	0.001	0.002
B	μ	-0.19	-0.04	-0.22	0.47	0.095	0.042	-0.002	-0.017
	σ_s	0.06	0.02	0.01	0.03	0.001	0.002	0.001	0.002

Table 1: Results of the experiments with two subjects obtained from runs over 30 minutes.

the parameters for the handle bar action are almost similar for the two subjects. However, the upper body actions differ considerably. For a better insight the Nyquist plots of the open loop transfer function were determined for both the subjects. For this purpose the block diagrams of the Figures 2 and 4 have been simplified to that of Fig. 7. Note that in Fig. 7 the signs at the subtraction point of the input φ_o have been reversed with regard to Fig. 4, and that therefore the transfer functions H_1 and H_2 have been provided with a minus sign. Taking into account that $\varphi_o = 0$ it follows from Fig. 7 that:

$$\Phi(\nu) = \frac{H_6(\nu)H_7(\nu)R(\nu) + H_3(\nu)H_7(\nu)N_1(\nu) - H_4(\nu)H_7(\nu)N_2(\nu)}{1 + H_7(\nu)\{H_2(\nu)H_4(\nu) - H_1(\nu)H_3(\nu) + H_5(\nu)\}}. \quad (19)$$

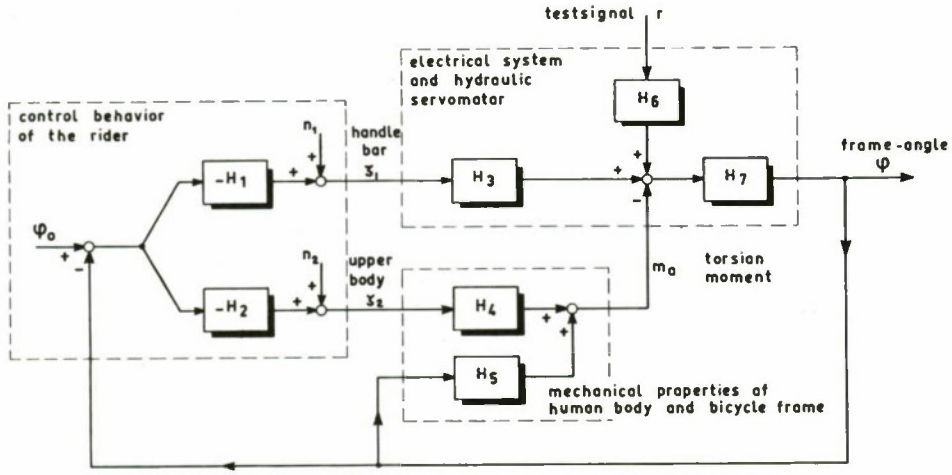


Figure 7: Block diagram of the man-bicycle system for a stabilization task.

Hence it follows for the open loop transfer function:

$$\overset{o}{H}(\nu) = H_7(\nu)\ \{H_2(\nu)H_4(\nu) - H_1(\nu)H_3(\nu) + H_5(\nu)\} \tag{20}$$

The calculation of $\overset{o}{H}(\nu)$ from Eq. 20 is not very accurate as the mechanical properties of the human body which dominate the transfer functions $H_4(\nu)$ and $H_5(\nu)$ are not known exactly. However, from Eq. 19, $\overset{o}{H}(\nu)$ can be derived as follows:

$$\frac{\Phi(\nu)}{R(\nu)} = H^*(\nu) = \frac{H_6(\nu)H_7(\nu)}{1 + \overset{o}{H}(\nu)}, \text{ or } \overset{o}{H}(\nu) = \frac{H_6(\nu)H_7(\nu)}{H^*(\nu)} - 1. \tag{21}$$

In this equation, $H_6(\nu)$ and $H_7(\nu)$ are well-known characteristics of the simulator, while $H^*(\nu)$ can be calculated from:

$$H^*(\nu) = \frac{S_{r\varphi}(\nu)}{S_{rr}(\nu)}\ . \tag{22}$$

The functions $S_{r\varphi}(\nu)$ and $S_{rr}(\nu)$ are the spectral densities belonging to the test signals r and the frame angle φ respectively. In this manner the data of the measurements for the two subjects were processed. The results are given in

Fig. 8. The figure indicates that there is quite a

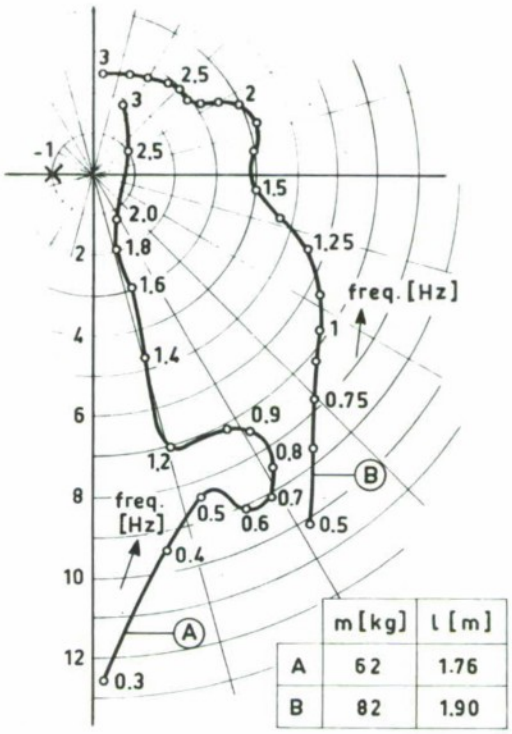


Figure 8: Nyquist plot for the open loop transfer function of the man-bicycle system of 2 subjects A and B.

difference between the control behavior of both the subjects. However, it should be noted that the weight and body dimensions of the subjects differ from each other. Furthermore, the plots pass the point -1 at a considerable distance to the right, which illustrates the fact that a trained subject has no difficulty to stabilize the simulator at a speed of 15 km/h. This agrees with the opinion of the subjects that they did not have to pay much attention to stabilize the bicycle. However, by making the stabilization task more critical the subjects might be forced to behave more uniformly. Similar experiments conducted by Jex [4] using a critical instability task, point towards this direction.

Finally, for a group of 5 male subjects C; D; E; F and G, characterized by Table 2, the parameters were determined for a total observation time of 10 minutes. For the sake of completeness, the data of the subjects A and B were added. From the values of the first and last 5 minutes the averaged values were obtained. The experiments were repeated each day for 10 days, between 2 p.m. and 3.30 p.m., under the

SUB-JECT	WEIGHT [kg]	LENGTH [m]	AGE [years]	MEAN AND VARIANCE OF RMS VALUES [°]			
				σ_φ	σ_{ζ_1}	σ_{ζ_1}	
A	62	1.76	33	0.69 0.04	0.84 0.02	0.56 0.04	μ σ
B	82	1.90	24	0.79 0.04	1.03 0.05	0.67 0.03	μ σ
C	66	1.78	20	0.62 0.09	0.91 0.18	0.97 0.10	μ σ
D	74	1.78	24	0.57 0.05	0.50 0.11	0.88 0.09	μ σ
E	63	1.63	23	0.75 0.11	0.57 0.10	0.72 0.06	μ σ
F	67	1.67	36	0.46 0.05	0.48 0.10	0.34 0.13	μ σ
G	71	1.83	26	0.71 0.15	0.72 0.07	0.95 0.10	μ σ

Table 2: Characteristics of the subjects observed and the RMS values of input and outputs of the bicycle rider.

same conditions. Each subject performed one experiment per day. The mean values μ and the standard deviations σ of the parameters over the 10 experiments for each subject are presented in Table 3. Moreover, the same quantities are given for the entire group in Table 4. The results obtained show that the integrative action for both handle bar and upper body can be neglected, hence the biased transfer functions based on the results of Table 4 become:

$$\left.\begin{aligned} \hat{H}_1(s) &= +1.07(1 + 0.15s)\, e^{-0.16s} \text{ (handle bar)} \\ \hat{H}_2(s) &= -0.13(1 + 1.6s)\, e^{-0.09s} \text{ (upper body)} \end{aligned}\right\} \qquad (23)$$

	SUB-JECT		PARAMETERS			REM-NANT	TIME DELAY	SENSITIVITIES		
			α_{11}	α_{12}	α_{13}	E_1	τ_1	$\frac{\Delta\alpha_{11}}{30\Delta\tau}$	$\frac{\Delta\alpha_{12}}{30\Delta\tau}$	$\frac{\Delta\alpha_{13}}{30\Delta\tau}$
HANDLE BAR CONTROL	C	μ	1.12	-0.003	0.17	0.31	0.161	-0.08	0.003	0.047
		σ	0.12	0.004	0.03	0.03	0.009	0.02	0.002	0.012
	D	μ	1.17	-0.008	0.14	0.30	0.159	-0.07	0.003	0.044
		σ	0.09	0.003	0.02	0.06	0.014	0.03	0.003	0.008
	E	μ	0.78	-0.000	0.15	0.22	0.146	-0.03	0.001	0.029
		σ	0.03	0.002	0.04	0.04	0.012	0.02	0.001	0.006
	F	μ	1.23	-0.001	0.21	0.32	0.156	-0.12	0.004	0.045
		σ	0.09	0.008	0.04	0.02	0.010	0.05	0.002	0.014
	G	μ	1.05	-0.006	0.14	0.29	0.171	-0.06	0.002	0.040
		σ	0.14	0.003	0.03	0.04	0.020	0.03	0.002	0.007
	SUB-JECT		α_{21}	α_{22}	α_{23}	E_2	τ_2	$\frac{\Delta\alpha_{21}}{30\Delta\tau}$	$\frac{\Delta\alpha_{22}}{30\Delta\tau}$	$\frac{\Delta\alpha_{23}}{30\Delta\tau}$
UPPER BODY CONTROL	C	μ	-0.17	-0.004	-0.23	0.62	0.096	0.11	-0.003	-0.014
		σ	0.18	0.017	0.01	0.04	0.004	0.03	0.001	0.008
	D	μ	+0.17	-0.021	-0.25	0.48	0.095	0.11	-0.003	-0.010
		σ	0.09	0.007	0.02	0.04	0.003	0.03	0.001	0.005
	E	μ	-0.46	-0.005	-0.19	0.36	0.090	0.03	-0.001	-0.019
		σ	0.17	0.004	0.03	0.07	0.006	0.01	0.0004	0.005
	F	μ	-0.18	-0.020	-0.16	0.57	0.089	0.09	-0.002	-0.013
		σ	0.12	0.004	0.01	0.07	0.008	0.03	0.0005	0.004
	G	μ	-0.03	-0.020	-0.24	0.48	0.091	0.08	-0.002	-0.010
		σ	0.12	0.010	0.02	0.05	0.004	0.04	0.001	0.005

Table 3: The mean values and standard deviations of the parameters, remnants, time delays and sensitivities for the five male subjects.

	α_{i1}	α_{i2}	α_{i3}	E_i	τ_i	$\frac{\Delta\alpha_{i1}}{30\Delta\tau}$	$\frac{\Delta\alpha_{i2}}{30\Delta\tau}$	$\frac{\Delta\alpha_{i3}}{30\Delta\tau}$
HANDLE BAR μ	1.07	-0.004	0.16	0.29	0.159	-0.07	0.003	0.041
CONTROL σ	0.19	0.005	0.05	0.05	0.016	0.04	0.002	0.012
UPPER BODY μ	-0.13	-0.014	-0.21	0.50	0.092	0.08	-0.002	-0.013
CONTROL σ	0.25	0.012	0.04	0.10	0.006	0.04	0.001	0.006

Table 4: The mean values and standard deviations of the parameters, remnants, time delays and sensitivities for a group of five male subjects.

Consecutively, the following standard deviations have been considered:

a. For the subjects A and B: Out of the standard deviations σ_5 given by Table 1, the mean standard deviations σ over two subjects were calculated. Under the assumption that the behavior of the subjects was time-invariant over a 10 minute run, the related standard deviations σ_a were derived from $\sigma_a = \sigma/\sqrt{2}$.

b. For the subjects C, D, E, F and G: From the standard deviations σ, which are given in Table 3 and which refer to 10 minute runs, the mean values σ_b over 5 subjects were determined. It should be noted that the standard deviations in Table 3 are obtained from 10 testruns performed on separate days.

c. For the group of subjects C, D, E, F and G as a whole: the standard deviations σ_c over all 50 measurements mentioned under item b were determined.

A comparison is made between the outcomes σ_a, σ_b and σ_c with reference to the 16 quantities. The results of this comparison are presented in Table 5 in the form of the ratios σ_a/σ_b and σ_c/σ_b. In the same way as indicated in Fig. 6, the mean values and dispersions of the ratios between the standard deviations for the 16 quantities were calculated.

σ_a/σ_b mean value: 0.54 dispersion: 0.33

σ_c/σ_b mean value: 1.67 dispersion: 0.74.

HANDLE BAR	α_{11}	α_{12}	α_{13}	E_1	τ_1	$\frac{\Delta\alpha_{11}}{30\Delta\tau}$	$\frac{\Delta\alpha_{12}}{30\Delta\tau}$	$\frac{\Delta\alpha_{13}}{30\Delta\tau}$
σ_a/σ_b	0.32	1.23	0.51	0.57	0.53	0.16	0.29	0.24
σ_c/σ_b	1.98	1.28	4.24	1.35	1.23	1.50	1.18	1.30
UPPER BODY	α_{21}	α_{22}	α_{23}	E_2	τ_2	$\frac{\Delta\alpha_{21}}{30\Delta\tau}$	$\frac{\Delta\alpha_{22}}{30\Delta\tau}$	$\frac{\Delta\alpha_{23}}{30\Delta\tau}$
σ_a/σ_b	0.16	1.34	0.46	0.66	0.67	0.57	0.72	0.27
σ_c/σ_b	1.82	1.43	2.23	1.92	1.15	1.51	1.43	1.15

Table 5: Ratios between the standard deviations obtained under different conditions.

Hence the estimated ratio between the three values of the standard deviations can be given by $\sigma_a : \sigma_b : \sigma_c = 1 : 2 : 3$.

6. Concluding remarks

Investigations on the behavior of a rider in a stabilization task have given the following results.

The dynamics of the bicycle rider can be described by a model consisting of a proportional-plus-derivative controller and a time delay. This model gives the relation between the input of the rider (the frame angle) and the outputs (the rotations of handle bar and upper body).

The behavior of the rider is time-invariant over at least a 5 minute period.

The bias effect resulting from the application of an open loop method in a closed loop system, is small for low values of the remnant. For higher values there is a strong influence on the gain parameter.

The standard deviations of the delay times are in the order of magnitude of 0.01 sec. An error in the delay time of 0.03 sec introduces an error in the estimated parameters which is of the same order as the standard deviations of these parameters. This implies that the procedure to determine the delay times is acceptable.

A Nyquist plot for 2 subjects shows that the stabilizing task in this particular investigation was subcritical, hence the subjects are rather free in choosing a control strategy.

If σ_a is the mean value of the standard deviation of the parameters for one subject within one test, if σ_b is the mean value of the standard deviation for one subject over a number of separate tests, and if σ_c is the mean value of the standard deviation for a group of subjects, then the relation between these quantities can be approximated by $\sigma_a : \sigma_b : \sigma_c = 1 : 2 : 3$.

7. Further research

So far, only the stabilization task has been studied. Recently the simulator has been extended with a visual display unit, so that a two degree of freedom task - stabilization and course following - can be investigated. (see Fig. 9).

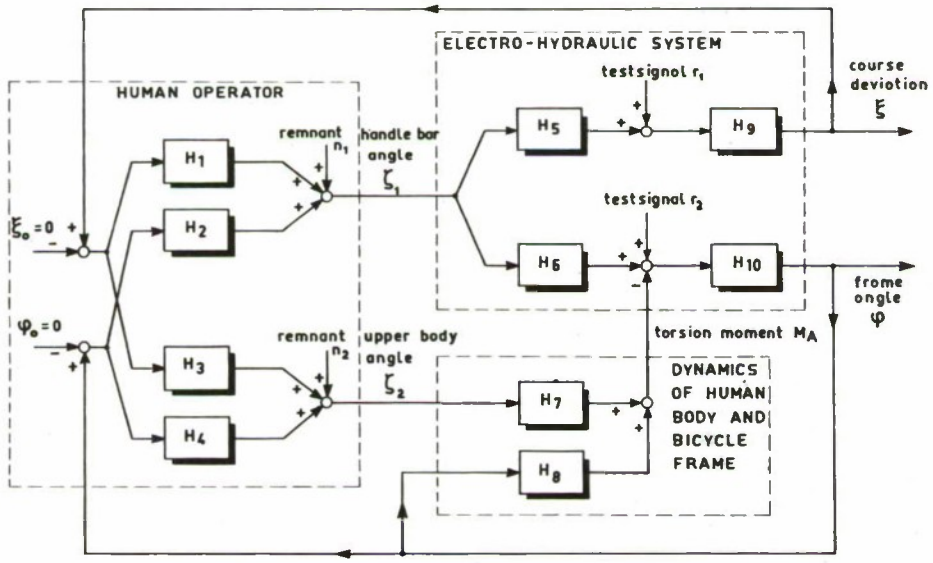


Figure 9: Block diagram of the man-bicycle system for a course following task and for a stabilization task.

In particular, the attention is focussed on a visual tracking task, where the stabilization of the simulator can be used as a critical-instability task [4].

The remnant originating from the rider presents a topic which requires further research. For instance, the question arises: Is the remnant caused by non-linearities, or should the remnant be considered as a test signal introduced by the rider? Preliminary investigations into the nature of input and outputs of the rider showed the occurence of energy peaks at frequencies of 0.5 and 2 Hz. The peak at 2 Hz may be due to the natural frequency of the system; the peak at 0.5 Hz was found when no external disturbances were introduced, so this might result from a test signal generated by the rider to get information about the system dynamics.

8. References

1. McRuer, D.T. et al.: New approaches to human-pilot/vehicle dynamic analysis, technical report AFFDL-TR-67-150, STI (1968), pp. 21-22.
2. Young, L.R.; Meiry, J.L.: Manual control of an unstable system with visual and motion cues, IEEE International Cong. Rec., vol. 13 pt. 6, (1965), pp. 123-127.
3. Donaldson, P.E.K.: Error decorrelation studies on a human operator performing a balancing task, Med. Electr. and Biol. Engng., vol. 2 No 4, (1964), pp. 393-410.
4. Jex, H.R.: Two applications of a critical-instability task to secondary work load research, IEEE Transactions on Human Factors in Electronics, vol. HFE 8 No 4, (1967), pp. 279-282.
5. Lunteren, A. van; Stassen, H.G.: Investigations on the characteristics of a human operator stabilizing a bicycle model, Intern. symp. on ergonomics in machine design, Prague (1967), p. 27.
6. Lunteren, A. van; Stassen, H.G.: On-line parameter estimation of the human transfer in a man-bicycle system, IVth IFAC Congres, paper 70.3, Warsaw (1969), p. 17.
7. Lunteren, A. van; Stassen, H.G.: Annual Report 1969 of the Man-Machine Systems Group, Laboratory for Measurement and Control, Dept. of Mech. Eng., Delft University of Technology, Delft (1970), Ch. II and III.
8. Whipple, F.J.W.: The stability of the motion of a bicycle, Quarterly Journal of Pure and Applied Math., vol. XXX, (1899), pp. 312-348.

9. Williams, M.; Lissner, H.R.: Biomechanics of human motion, W.B. Saunders Cy, London (1962), pp. 134-135.
10. Stassen, H.G.: The polarity coincidence correlation technique - A usefool tool in the analysis of human-operator dynamics, IEEE transactions on man-machine systems, vol. MMS-10 No 1, (1969), pp. 34-39.
11. McRuer, D.T.; Krendel, E.S.: A review and summary of tracking research applied to the description of human dynamic response, Wescon Conv. Rec. (1958), pp. 254-262.
12. Tustin, A.: The nature of the operator's response in manual control and its implications for controller design, Journal of the IEEE, vol. 94. pt. IIA, (1947), pp. 190-202.
13. Ragazzini, J.R.: Engineering aspects of the human being as a servomechanism. Unpublished paper presented at the Am. Psychol. Assn. Meeting (1948).
14. Booton, R.C.: The analysis of non-linear control systems with random inputs, Symp. on non-linear circuit analysis, Polytechn. Inst. of Brooklyn (1953), pp. 369-391.
15. McRuer, D.T.; Graham, D.: Pilot-vehicle control system analysis, Progress in astronautics and aeronautics, vol. 13, (1964), pp. 603-621.
16. Fogel, L.J.: An analysis for human flight control, IRE-Convention rec. pt. 8, (1956), pp. 69-88.
17. Elkind, J.L.: Characteristics of simple manual control systems, Techn. Report No 111, MIT, Lincoln Lab. (1956).

ON THE OPTIMALITY OF THE HUMAN OPERATOR

Glenn A. Jackson

Assistant Professor of Engineering

Oakland University

Rochester, Michigan 48063

Abstract

A study has been conducted on the optimal properties of the human operator in low order compensatory systems. The study was limited to those cases where the approximate McRuer crossover model

$$\frac{\theta_o}{e}(s) = \frac{K(\frac{2}{\tau} - s)}{s(\frac{2}{\tau} + s)}$$

is applicable. It was shown that, for the subjects tested, the parameters K and τ were effectively adjusted to minimize the cost functional

$$J(K,\tau) = \overline{e^2(t,K,\tau)} + \alpha\overline{c^2(t,K,\tau)}$$

with the constraint

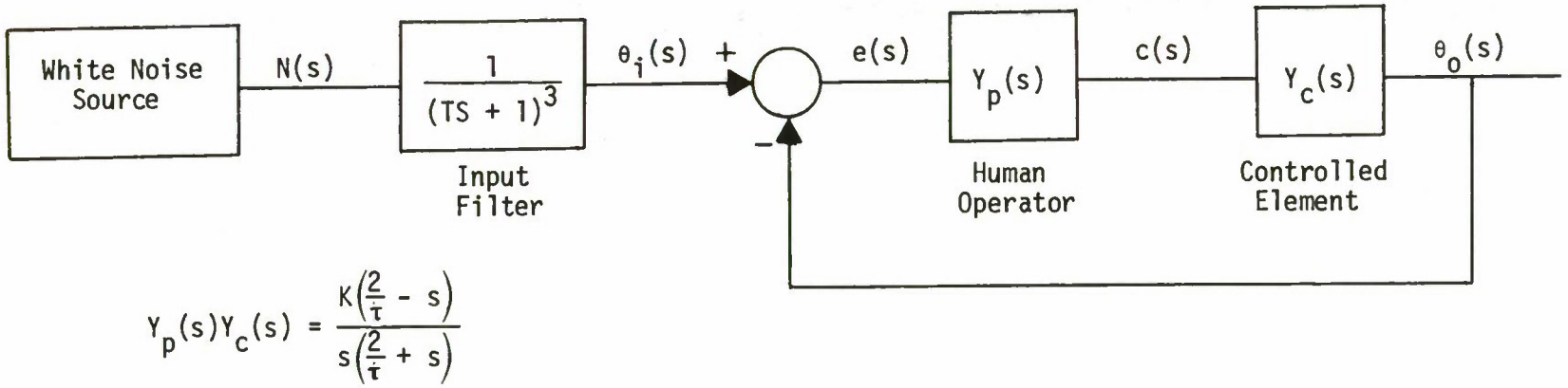
$$\tau = \beta K,\ 0.01 \leq \beta \leq 0.12 \text{ seconds}^2.$$

$\overline{e^2}$ is the mean squared value of the compensatory system error and $\overline{c^2}$ is the mean squared value of control effort. β is inversely proportional to the square of the subject's closed loop bandwidth, and thus varies between subjects, and even within the same subject at various stages of training. For a given controlled element α is a constant that doesn't change either between subjects or with subject practice. However, α does change with the type of system being controlled and becomes smaller as the control task becomes more difficult.

The conclusion is drawn that all subjects optimize the same cost functional when given the same control task. They perform differently, however, due to their varying abilities in processing high frequency signals.

A. Introduction

In a series of tests previously reported by the author [1,2], the parameters K and τ of an approximate crossover model of the compensatory control system were measured directly using parameter tracking techniques. The testing was accomplished using six subjects: three controlling a 5/s element and three controlling a $5/s^2$ element. In each case input filter time constants of 1, 1/2 and 1/4 seconds were used, and K and τ values were determined for each subject during each day of training. The assumed model for the tests conducted is given in Figure 1.



$$Y_p(s)Y_c(s) = \frac{K\left(\frac{2}{\tau} - s\right)}{s\left(\frac{2}{\tau} + s\right)}$$

Figure 1. Compensatory System Model with Input Signal Generation.

One important property of the data generated by these tests, which was not discussed in the previous analysis, is that in the cases where the approximate crossover model was exceptionally good*, the K and τ values were highly correlated. This correlation existed at a given test condition, between K and τ values of different subjects, and even between the values for the same subject that were determined on different days during the training period.

In Tables 1 and 2 in the Appendix the daily average values of K and τ are given for each subject at each of the test conditions where the approximate crossover model was considered to be exceptionally good. These test conditions were with controlled elements, $Y_c(s)$, and input filter time constants, T, of:

(1) $Y_c(s) = 5/s$, T = 1 second,

(2) $Y_c(s) = 5/s$, T = 1/2 second,

(3) $Y_c(s) = 5/s^2$, T = 1 second.

In Figure 2 the data points obtained from these conditions are plotted along with the linear regression line for each condition. The correlation coefficients are also noted in the figures.

In the remainder of this paper it will be shown that the correlation existing between the measured values of K and τ can be attributed to an apparent parameter optimization of the approximate crossover model. It will be shown that when the K and τ parameters of the model are adjusted to minimize a realistic performance index at a given test condition, the optimum values fall on essentially the same linear regression lines as found in Figure 2.

B. Development of an Optimal Policy

To account for the correlation existing between the experimental values of K and τ, three postulates are developed.

Rationale for Postulate 1: At any stage of training, each subject has the capability of operating the closed loop compensatory system at some maximum bandwidth. Further, for the low input frequencies being discussed, the closed loop bandwidth can be approximated by the natural frequency of the compensatory system. This approximation is reasonable since the input power at frequencies above the natural frequency can be shown to be extremely low [1].

In terms of the approximate crossover model the closed loop natural frequency, ω_n, is

$$\omega_n = \sqrt{\frac{2K}{\tau}} \ .$$

*"Exceptionally good" means that the power in the error between model output and compensatory system output is less than 10% of the power out of the compensatory system. The Power Match [1] is thus greater than 90%. This indicates that the model is accurate and that the human operator is fairly linear.

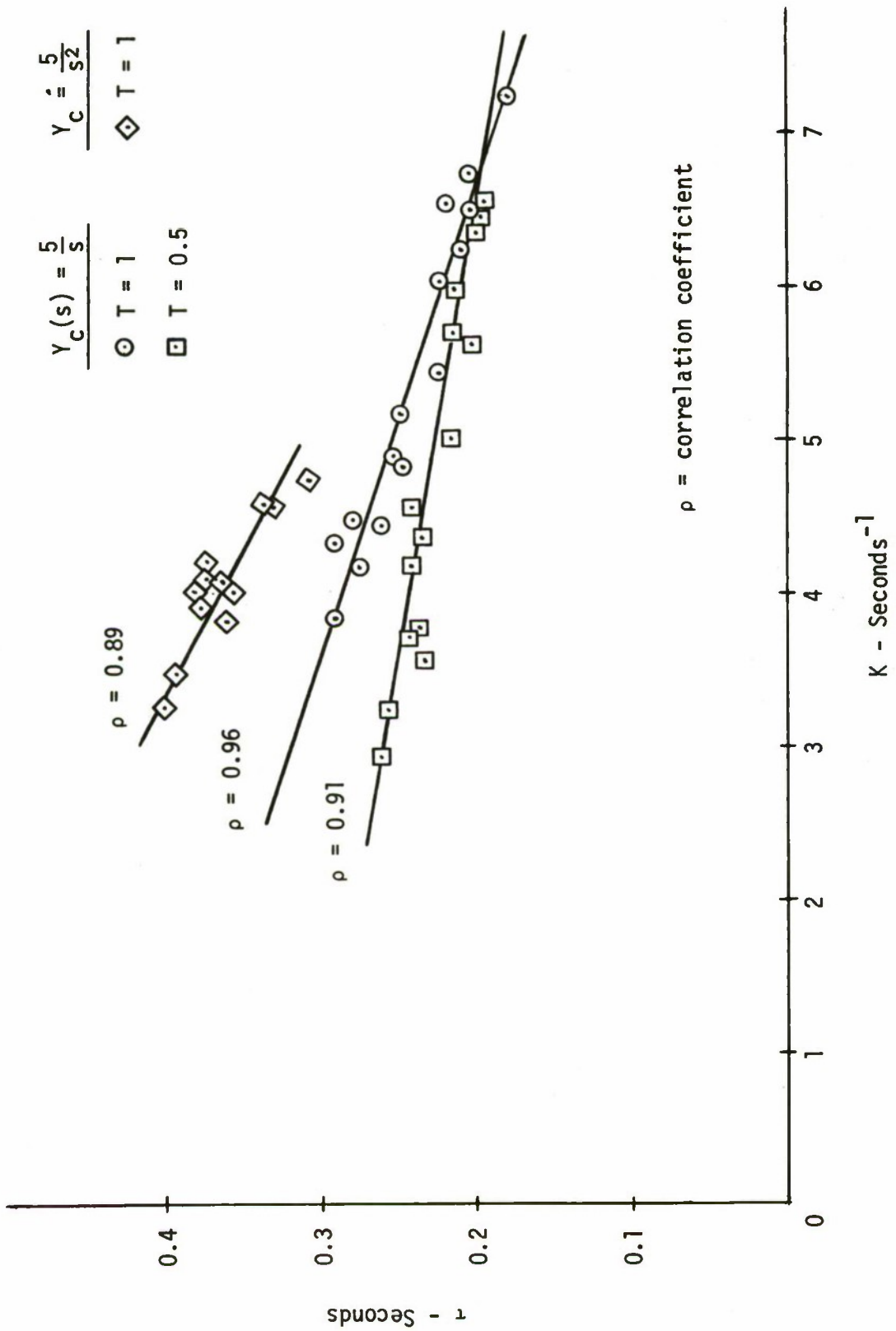


Figure 2. Correlation of the Parameters K and τ

If it is assumed that each subject operates the compensatory system at his maximum bandwidth, and that this maximum is fixed during a given interval of his training, then

$$\sqrt{\frac{2K}{\tau}} = \text{constant} = B = \text{Bandwidth in radians/second.}$$

$$\frac{2K}{\tau} = B^2$$

$$\tau = \frac{2}{B^2} K$$

or

$$\tau = \beta K. \tag{1}$$

Under the conditions assumed, Equation (1) can be considered as a constraint on any optimal adjustment of the parameters K and τ which might be taking place. It should be emphasized that β in Equation (1) will be different for different subjects, and different for the same subject during various intervals of his training period. Emperically, β can be determined to be in the range $0.01 \leq \beta \leq 0.12$ seconds^2.

Postulate 1: Any parameter optimization of the approximate crossover model (compensatory control system) is constrained by Equation (1): $\tau = \beta K$, $0.01 \leq \beta \leq 0.12$. β is fixed for a given subject, at a given point in training, on a given tracking task.

Rationale for Postulate 2: A general performance index in optimal control theory, and one that seems intuitively reasonable for the human operator, involves the combination of control system error and control effort. One form of this index is

$$J(K,\tau) = E\left\{\frac{1}{t_f}\int_0^{t_f}\left[e^2(t,K,\tau) + \alpha c^2(t,K,\tau)\ dt\right]\right\}$$

or

$$J(K,\tau) = \overline{e^2} + \alpha\overline{c^2} \tag{2}$$

where $E\{\cdot\}$ is the statistical average, t_f is the time of the tracking task, e^2 is the mean squared value of the tracking error, and c^2 is the mean squared value of the signal out of the control stick. α is a constant.

Postulate 2: The performance index being minimized by the human operator is Equation (2): $J(K,\tau) = e^2 + \alpha c^2$.

Postulate 3: The control strategy of the human operator, while controlling one of the elements under discussion, is a two-step process consisting of:

(1) Forcing the compensatory system into the form of the approximate crossover model.

(2) Minimizing the performance index

$$J(K,\tau) = \overline{e^2} + \alpha\overline{c^2}$$

via a parameter adjustment that is constrained by the relation $\tau = \beta K$.

Since part (1) of Postulate 3 has already been verified by several researchers [3], it remains only to study the feasibility of part (2).

C. Analysis of the Postulated Strategy

To evaluate the optimal strategy postulated above, and to determine the value of α for each test condition, a digital computer program was developed to generate J for the approximate crossover model. This was done for the three test conditions noted above. For

Condition 1: $Y_c(s) = 5/s$ and $T = 1$ second, and

Condition 2: $Y_c(s) = 5/s$ and $T = 1/2$ second,

the transfer function relating c(s) to e(s) (i.e., the transfer function for the human operator) was assumed to be

$$\frac{c}{e}(s) = \frac{K(2/\tau - s)}{5(2/\tau + s)}$$

so that the entire forward loop remained as the approximate crossover model. For

Condition 3: $Y_c(s) = 5/s^2$ and $T = 1$ second,

$$\frac{c}{e}(s) = \frac{Ks(2/\tau - s)}{5(2/\tau + s)}$$

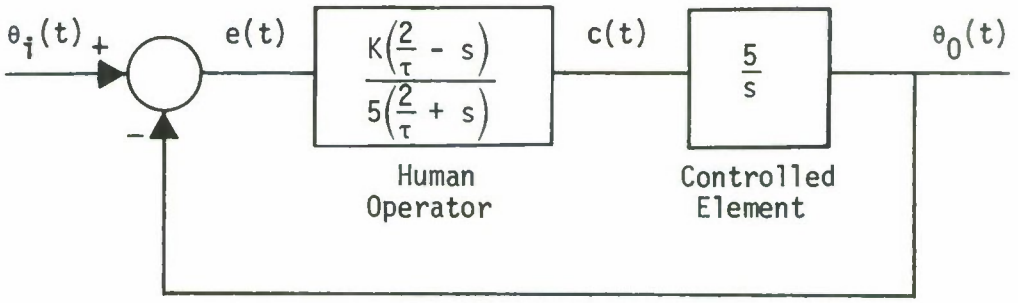
so that the approximate crossover model was retained. Figures 3a and 3b show the compensatory system models for the three conditions.

In terms of the block diagrams of Figure 3, the performance index can be reduced to the following integral forms.

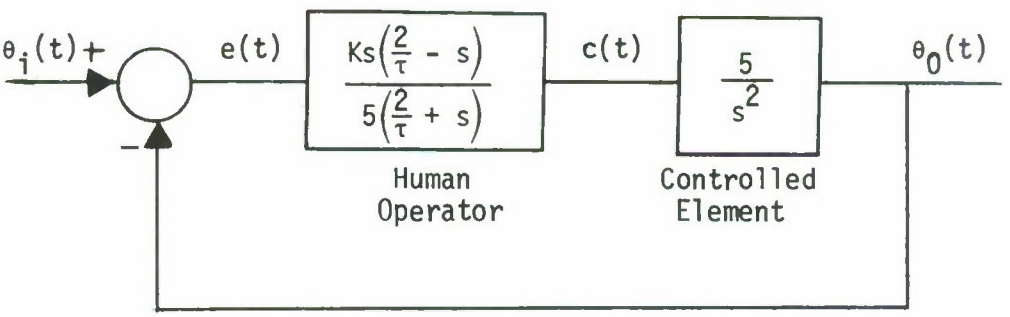
$$J(K,\tau) = \overline{e^2} + \alpha\overline{c^2}$$

$$= \frac{1}{2\pi j}\left\{ \int_{-j\infty}^{j\infty} \left|\frac{\theta_i}{N}(s)\right|^2 \left|\frac{e}{\theta_i}(s)\right|^2 \Phi_{NN}(s)\, ds \right.$$

$$\left. + \alpha \int_{-j\infty}^{j\infty} \left|\frac{\theta_i}{N}(s)\right|^2 \left|\frac{c}{\theta_i}(s)\right|^2 \Phi_{NN}(s)\, ds \right\} .$$



(a) $Y_c(s) = \frac{5}{s}$



(b) $Y_c(s) = \frac{5}{s^2}$

Figure 3. Compensatory System Models for Two Controlled Elements Based on Approximate Crossover Model

$\Phi_{NN}(j\omega)$ is the power spectral density of the white noise source, N(t). (See Figure 1). Assuming that $\Phi_{NN}(j\omega) = 1$,

$$J(K,\tau) = \frac{1}{2\pi j}\left\{ \int_{-j\infty}^{j\infty} \left|\frac{1}{(Ts+1)^3}\right|^2 \cdot \left|\frac{s(2/\tau + s)}{s^2 + (2/\tau - K)s + 2K/\tau}\right|^2 ds \right.$$

$$\left. + \alpha \cdot \int_{-j\infty}^{j\infty} \left|\frac{1}{(Ts+1)^3}\right|^2 \cdot \left|\frac{(Ks/5)(2/\tau - s)}{s^2 + (2/\tau - K)s + 2K/\tau}\right|^2 ds \right\} \qquad (3)$$

for Conditions 1 and 2, and

$$J(K,\tau) = \frac{1}{2\pi j}\left\{ \int_{-j\infty}^{j\infty} \left|\frac{1}{(Ts+1)^3}\right|^2 \cdot \left|\frac{s(2/\tau + s)}{s^2 + (2/\tau - K)s + 2K/\tau}\right|^2 ds \right.$$

$$\left. + \alpha \cdot \int_{-j\infty}^{j\infty} \left|\frac{1}{(Ts+1)^3}\right|^2 \cdot \left|\frac{(Ks^2/5)(2/\tau - s)}{s^2 + (2/\tau - K)s + 2K/\tau}\right|^2 ds \right\} \qquad (4)$$

for Condition 3. These equations were also normalized by dividing by $\overline{e^2}$ at K = 0. The normalization factor happens to be 3/16T in all cases. The solution of these equations was accomplished through the use of tabulated solutions of Parseval's Theorem [4], under the constraint, $\tau = \beta K$.

A programmed search of the solutions to Equations (3) and (4) was made in the following manner. With α, β and T fixed, $\min_{K,\tau} J$ was determined. This was done for various values of α, β and T in order to find that value of α which, for a given $Y_c(s)$ and T, would yield K and τ values equivalent to those present on the subject regression lines in Figure 3. If the postulated strategy is correct, and if all subjects are minimizing the same functional at a given test condition, then the optimum values of K and τ should fall on the regression lines of Figure 3, regardless of the value to which β is fixed. This was indeed shown to be the case.

The results of the computer analysis may be summarized as follows:

For Condition 1: $\alpha = 2.5$.

The $\min_{K,\tau} J$ with $\alpha = 2.5$ and $\tau = \beta K$, where β is any fixed value $0.01 \leq \beta \leq 0.08$, occurs at values of K and τ which fall essentially on the subject's $Y_c(s) = 5/s$, T = 1 regression line of Figure 3.

For Condition 2: $\alpha = 0.25$.

The $\min_{K,\tau} J$ with $\alpha = 0.25$ and $\tau = \beta K$, where β is any fixed value $0.01 \leq \beta \leq 0.08$, occurs at values of K and τ which fall essentially on the subject's $Y_c(s) = 5/s$, T = 1/2 regression line of Figure 3.

For Condition 3: $\alpha = 0$.

The $\min_{K,\tau} J$ with $\alpha = 0$ and $\tau = \beta K$, where β is any fixed value $0.07 \leq \beta \leq 0.12$, occurs at values of K and τ which fall essentially on the subject's $Y_c(s) = 5/s^2$, T = 1 regression line of Figure 3.

The optimum values of K and τ for several values of β are given in Figures 4, 5 and 6 for Conditions 1, 2 and 3, respectively. These figures indicate the strong relationship between the K-τ regression lines of Figure 3 and the optimum K-τ values derived from the postulated performance index.

D. Conclusions and Extensions

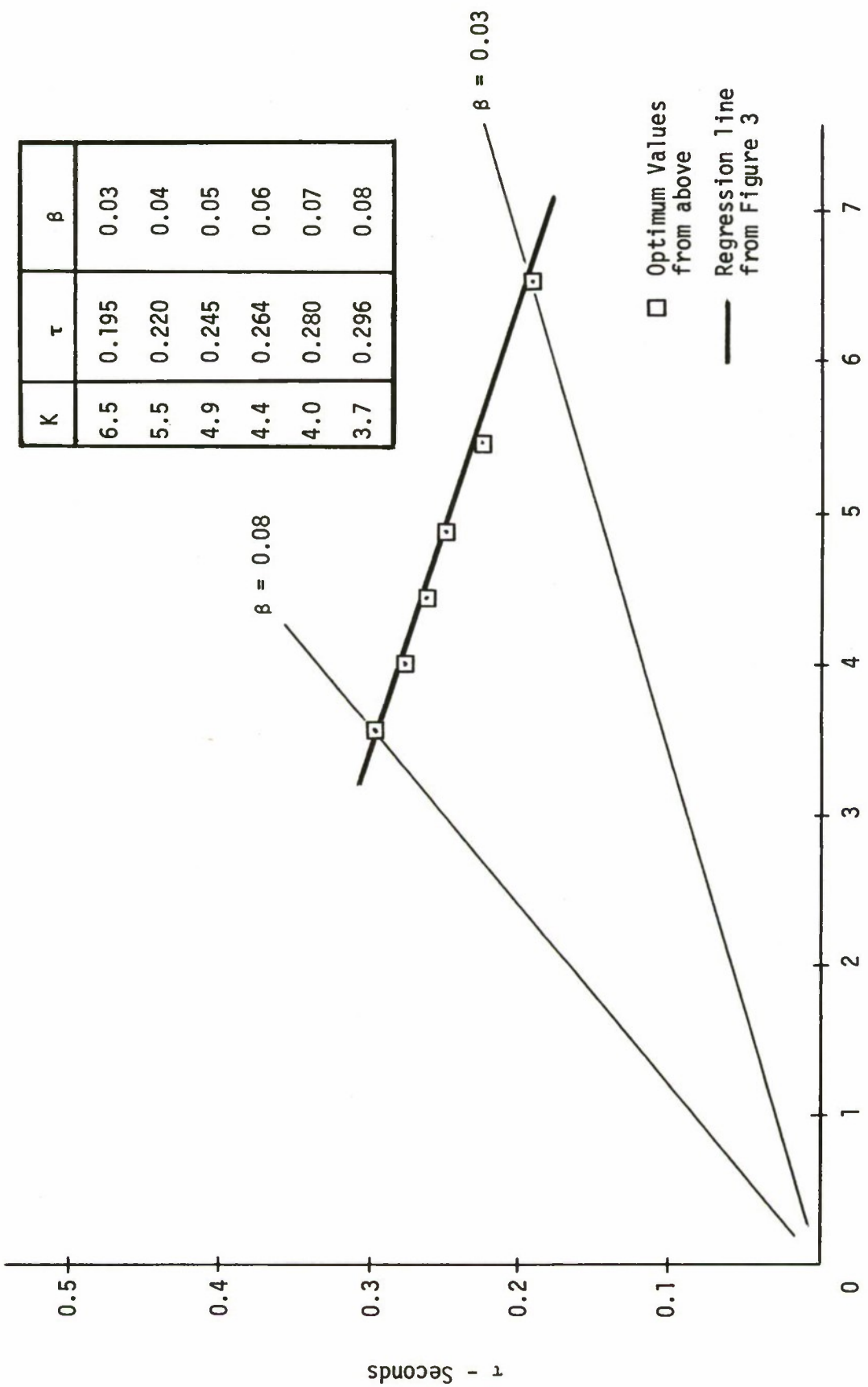
Conclusions:

From the data summarized in Figures 4-6 several tentative conclusions can be drawn.

(1) The postulated optimum strategy yields crossover model parameters which are in close agreement with the data obtained from six subjects and two controlled elements which were available for analysis. Thus, the hypothesis that subjects in compensatory control systems essentially minimize the performance index $J = e^2 + \alpha c^2$ for their associated crossover model, under the constraint $\tau = \beta K$, appears valid and warrants further experimental investigation.

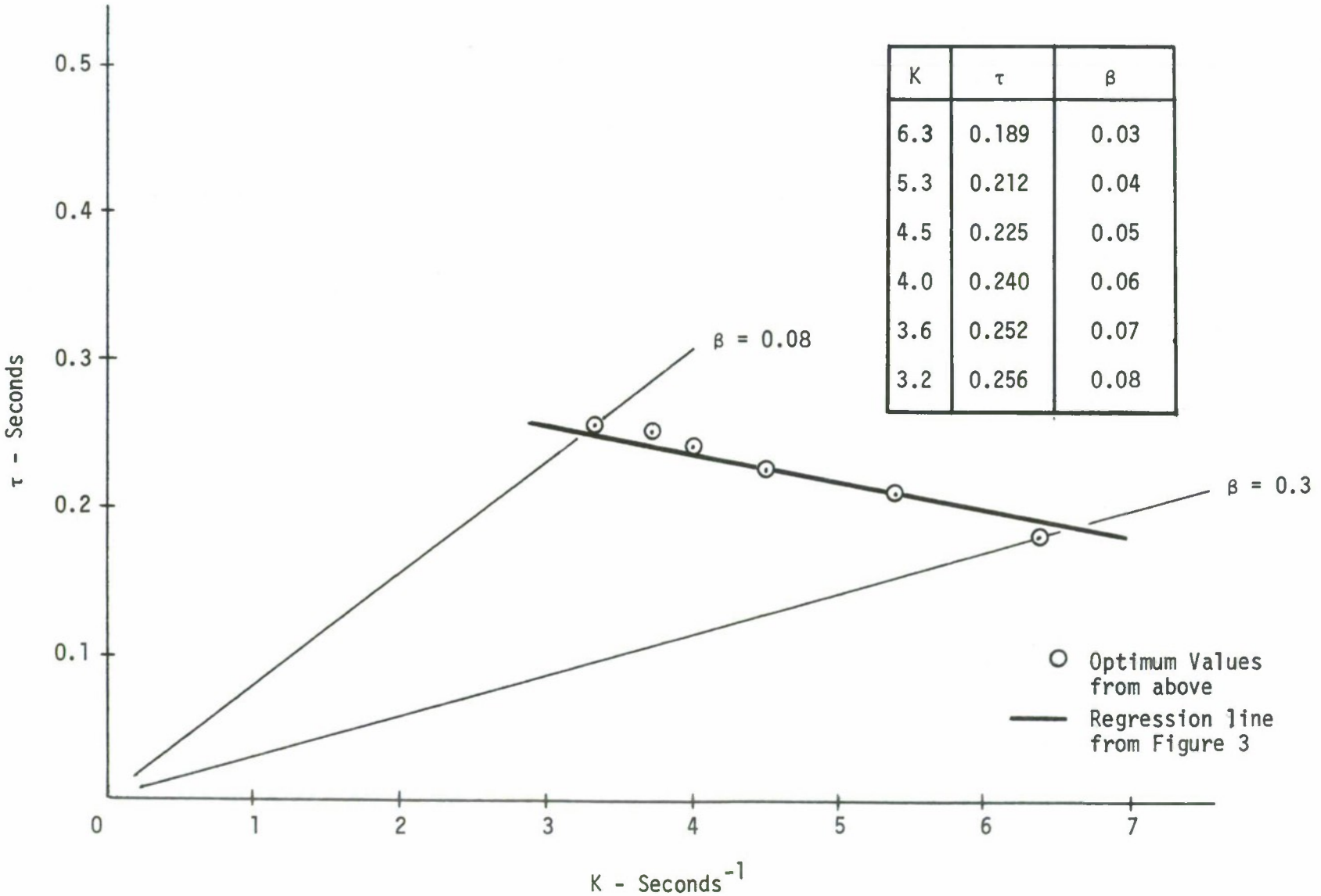
(2) For a given test condition a unique value of α in the performance index $J = e^2 + \alpha c^2$ exists which is the same for all subjects. The β's in the constraint $\tau = \beta K$, however, vary between subjects. The conclusion can be drawn that all subjects minimize the same performance index, but arrive at different parameter values due to the subject-to-subject variation in ability to control high frequency components.

(3) Test Conditions 1, 2 and 3 are arranged in order of increasing task difficulty. Since the value of α shows a definite tendency to decrease as task difficulty increases, the implication is that subjects penalize control effort more on easier tasks than they do for harder tasks. They can "keep the error small" and still conserve control energy when the task is easy. However, as the task becomes harder this is not true, and they thus trade off control effort in favor of reducing the system error, which is the assigned task.



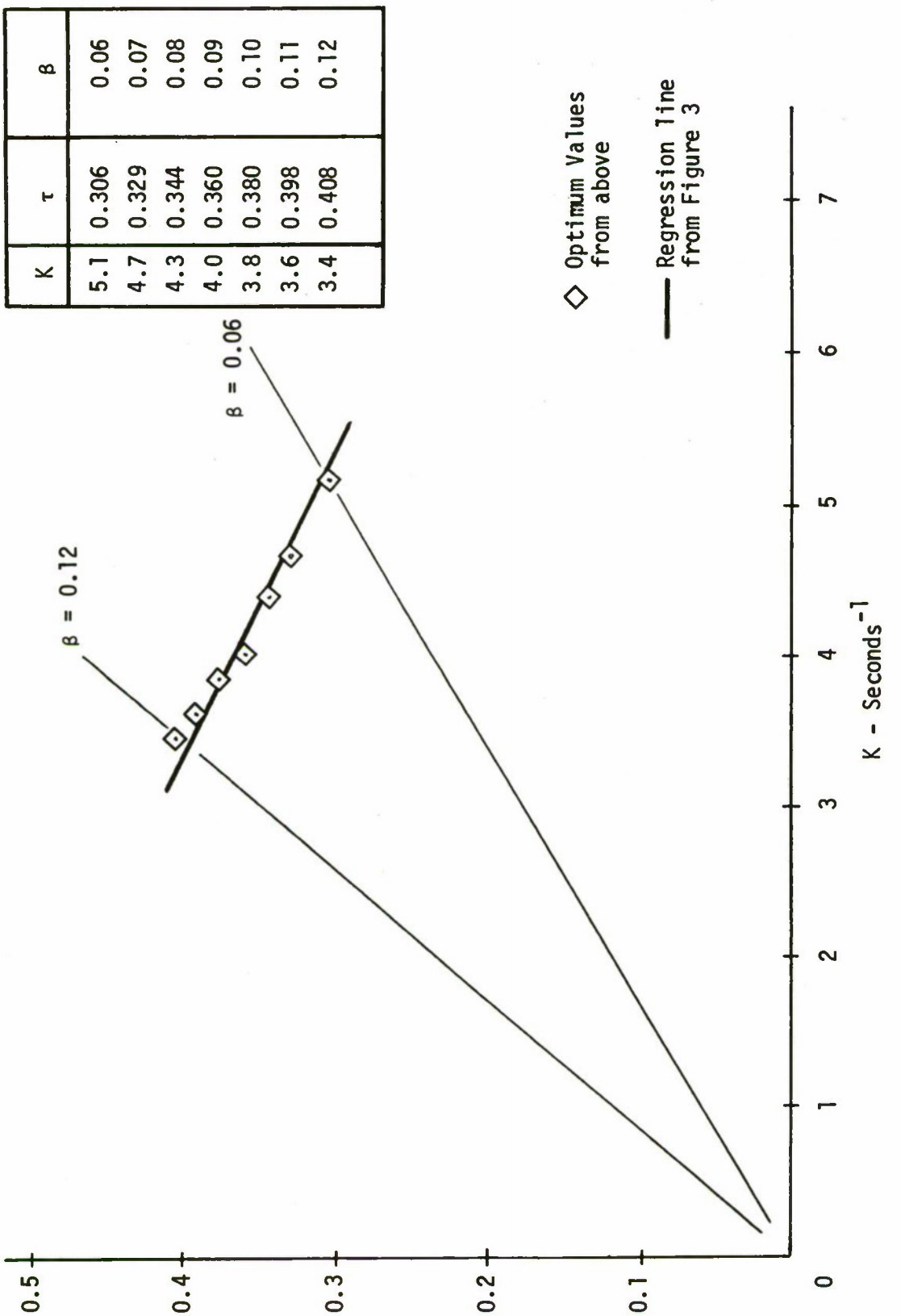
K	τ	β
6.5	0.195	0.03
5.5	0.220	0.04
4.9	0.245	0.05
4.4	0.264	0.06
4.0	0.280	0.07
3.7	0.296	0.08

Figure 4. Optimum K and τ Values: $J = \overline{e^2} + 2.5\ \overline{c^2}$, $Y_c(s) = \frac{5}{s}$, T = 1



K	τ	β
6.3	0.189	0.03
5.3	0.212	0.04
4.5	0.225	0.05
4.0	0.240	0.06
3.6	0.252	0.07
3.2	0.256	0.08

Figure 5. Optimum K and τ Values: $J = \overline{e^2} + 0.25\ \overline{c^2}$, $Y_c(s) = \frac{5}{s}$, T = 0.5



K	τ	β
5.1	0.306	0.06
4.7	0.329	0.07
4.3	0.344	0.08
4.0	0.360	0.09
3.8	0.380	0.10
3.6	0.398	0.11
3.4	0.408	0.12

Figure-6. Optimum K and τ Values: $J = e^{\overline{2}}$, $Y_c(s) = \frac{5}{s^2}$, T = 1

Extensions:

(1) Since only six subjects were used in obtaining the experimental data from which the postulates were drawn and evaluated, several additional subjects should be run to see if their model parameters fall on the same regression lines as found for the first subjects. This extension has been started and will be reported on at a later date.

(2) If the postulates are correct, the changes in the values of K and τ for a given trained subject, due to small variations in the spectrum of the input signal, should be predictable. An interesting area of research would be to see if this prediction can indeed be made.

E. Acknowledgment

This research was conducted under National Science Foundation Grant GK 3490.

F. References

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

2. Jackson, G. A.: A Method for the Direct Measurement of Crossover Model Parameters. IEEE Transactions on Man-Machine Systems, Vol. MMS-10, No. 1, pp. 27-33, March 1969.

3. McRuer, D. T., Graham, D., Krendel, E. S., and Reisener, W., Jr.: Human Pilot Dynamics in Compensatory Systems. AFFDL-TR-65-15, July 1965.

4. Eveleigh, V. W.: Adaptive Control and Optimization Techniques. McGraw-Hill Book Company, Appendix A, 1967.

G. Appendix

Each K and τ value in the following tables was determined by averaging over those K-τ values determined in 5 two-minute tracking tasks. These experimental tests are discussed at length in references [8, 9].

Table 1

$Y_c(s) = 5/s$

Subject No.	Day of Training	Input Filter Time Constant (Seconds)	K (seconds^{-1})	τ (seconds)
1	2	1.0	4.92	0.240
↓	2	0.5	3.63	0.221
	4	1.0	5.23	0.240
	4	0.5	4.61	0.234
	6	1.0	5.52	0.215

Table 1
(contd.)

Subject No.	Day of Training	Input Filter Time Constant (Seconds)	K ($seconds^{-1}$)	τ (seconds)
1	6	0.5	5.08	0.201
↓	8	1.0	6.55	0.194
	8	0.5	5.70	0.198
	10	1.0	6.79	0.198
	10	0.5	6.05	0.206
2	2	1.0	4.95	0.245
↓	2	0.5	4.24	0.231
	4	1.0	6.10	0.216
	4	0.5	5.75	0.206
	6	1.0	6.61	0.210
	6	0.5	6.30	0.205
	8	1.0	7.28	0.172
	8	0.5	6.39	0.192
	10	1.0	6.31	0.202
	10	0.5	6.62	0.188
3	2	1.0	3.93	0.282
↓	2	0.5	3.01	0.249
	4	1.0	4.42	0.281
	4	0.5	3.33	0.246
	6	1.0	4.25	0.266
	6	0.5	3.75	0.233
	8	1.0	4.50	0.253
	8	0.5	4.45	0.234
	10	1.0	4.54	0.273
	10	0.5	3.83	0.229

Table 2

$Y_c(s) = 5/s^2$

Subject No.	Day of Training	Input Filter Time Constant (Seconds)	K ($seconds^{-1}$)	τ (seconds)
4	3	1.0	3.32	0.392
↓	5		4.12	0.347
	7		3.90	0.353
	9		4.80	0.301
5	3	1.0	4.19	0.365
↓	5		4.30	0.367
	7		4.14	0.357
	9		4.65	0.329
6	3	1.0	3.56	0.384
↓	5		4.03	0.369
	7		4.10	0.373
	9		4.64	0.325

VISUAL SOURCES OF CONTROLLER REMNANT

By

William H. Levison

Bolt Beranek and Newman Inc.
Cambridge, Massachusetts

EXPANDED SUMMARY

The results of a current investigation of controller remnant are presented. The primary objective of this study is to investigate visual sources of remnant. Experiments have been conducted to explore the effects of display signal amplitude and display signal bandwidth on the equivalent observation noise spectrum.

Previous studies of controller remnant have led to the modelling of controller remnant in terms of an equivalent observation noise source; that is, each input variable utilized by the controller is assumed to be perturbed by a white noise process internal to the controller (Ref. 1,2). Since the pilot can theoretically obtain both rate and position information from a single display indicator, observation noise must be considered as a vector process. The following model structure has been considered:

$$\underline{R} = \underline{Q} + \underline{P} \cdot \underline{\sigma}^2 = \underline{P} \cdot (\underline{\sigma}_T^2 + \underline{\sigma}^2) \qquad (1)$$

where $\underline{R}$ is the injected white noise vector; $\underline{P}$ is a diagonalized matrix of normative white noise processes, or "noise ratios"; $\underline{\sigma}^2$ is the vector composed of the variances of the displayed signals; and $\underline{Q}$ and $\underline{\sigma}_T^2$ are noise processes to account for the irreducible, or threshold, component of observation noise.

The first experiment was designed to verify the model structure; i.e., to show that the power density level of each component noise process varied linearly with signal variance. Measurements were obtained from a set of single-axis tracking experiments in which the subjects were provided with a compensatory display of system error. Vehicle dynamics were K/s, and a simulated first-order noise spectrum (constructed from sums of sinusoids) was applied in parallel with the pilot's output to perturb the system. The control gain was chosen to minimize the effects of saturation and threshold effects associated with the neuromotor system. In order to achieve the desired varation in the variance of the displayed signal, the display gain was varied over a range of 20:1. The remaining experimental parameters were held fixed. Four different display gains were investigated under both foveal and peripheral viewing conditions.

A linear transformation was performed on the pilot's remnant spectrum to obtain an equivalent *scalar* observation noise process. (See reference 2 for a description of measurement techniques.) A model-matching procedure using the optimal-control model of Kleinman et al (Ref. 3) was then employed to resolve this scalar spectrum into the underlying *vector* process so that separate estimates could be obtained for the noise processes associated with the estimation of indicator position and indicator velocity.

Four plots of noise power density versus signal variance were obtained: one for position noise vs. display error variance and one for rate noise vs. display error rate variance for each of the two viewing conditions. For the most part, these relationships were highly linear, thus confirming the model structure of equation (1). The noise ratio $P_{(.)}$ was defined as the slope of

the noise-versus-variance curve, and the threshold noise $Q_{(.)}$ was obtained by extrapolating this curve to zero variance. The threshold levels showed reasonable trends; i.e., all estimates were positive numerically (negative values would have no physical meaning), and peripheral thresholds were consistently greater than corresponding foveal threshold terms. The four noise ratios obtained from these plots differed. The ratios associated with rate were greater than the ratios associated with position, and peripheral noise ratios were somewhat greater than corresponding foveal ratios.

Since one of the usual goals of model-building is to construct a model with the fewest number of parameters, we wished to determine whether or not we could justifiably replace the matrix noise ratio $\underline{P}$ by a single scalar P. That is, we investigate the modelling errors resulting from the assumption that all values of the $\underline{P}$ matrix were numerically identical. Accordingly, predictions of system performance were obtained via the optimal-control model using the observation noise model of equation (1) with all values of $\underline{P}$ set equal to the average of the four noise ratio measurements. The $\underline{Q}$ matrix was unchanged. Predicted system performance measures (system error and error rate variance scores, as well as variance scores for pilot control force and control rate) obtained from this model were compared with scores obtained from the model in which the four "best" noise ratios were used. Since corresponding scores obtained by these two modelling procedures differed in general by less than 10%, we conclude that little predictive accuracy is sacrificed by the assumption that the noise ratio is identical for all observations performed by the controller.

A second experiment was performed to determine whether or not the model structure of equation (1) would have to be modified to include the effects of signal bandwidth. The control situation

was similar to that described above, except that a second-order Butterworth filter was cascaded with the K/s dynamics to reduce signal bandwidth. Two cutoff frequencies (1.0 rad/sec and 2.0 rad/sec) were considered, and both foveal and peripheral viewing conditions were again investigated.

The results of this experiment are currently being analyzed. Since the subjects appear to have used acceleration information along with position and rate information to control the system, an acceleration component has been added to the observation noise vector to obtain a good match to the measured noise spectrum. We have not found consistent effects of bandwidth on the components of the observation noise vector, and the fixed-noise-ratio model appears to account for the effects of controller remnant on system performance, at least for foveal viewing.

In conclusion, the effects of controller remnant on system performance can be reasonably well accounted for by an equivalent observation noise model of the form

$$\underline{R} = \underline{Q} + P \cdot \underline{\sigma}^2 \qquad (2)$$

where the noise ratio P is assumed to be the same for all components of the vector process. The $\underline{Q}$ matrix accounts for irreducible noise terms which, for the experimental conditions that were investigated, can be attributed primarily to visual noise sources. The observation noise vector contains a component process to be associated with each of the display variables utilized by the controller. The specific set of variables that are useful to the controller will depend on the nature of the display, the vehicle dynamics, and, to some extent, on the disturbance function.

REFERENCES

1. Levison, W.H., Baron, S., and Kleinman, D.L., "A Model for Human Controller Remnant," IEEE Trans. on Man-Machine Systems, MMS-10:101-108, December 1969.

2. Levison, W.H., Elkind, J.I., and Ward, J.L., "Studies of Multivariable Manual Control Systems: A Model for Task Interference," Rep. No. 1892, (prepared for NASA/Ames Research Center under Contract No. NAS2-3080) Bolt Beranek and Newman Inc., Cambridge, Mass., December 8, 1969.

3. Kleinman, D.L., Baron, S., and Levison, W.H., "An Optimal Control Model of Human Behavior," Proceedings of the Fifth Annual NASA/University Conference on Manual Control, M.I.T., Cambridge, Mass., March 27-29, 1969.

RESEARCH ON A NEW HUMAN DYNAMIC RESPONSE TEST BATTERY*

Part I. Test Development and Validation
Part II. Psychophysiological Correlates

H. R. Jex and R. W. Allen
Systems Technology, Inc.
Hawthorne, Calif.

ABSTRACTS

PART I, TEST DEVELOPMENT AND VALIDATION

A battery of autopaced critical-instability tasks, subcritical tracking tasks, and step reaction-time tests was developed to permit efficient measurement of the limiting human dynamic response properties. Standard test parameters for first-, second-, and third-order controlled elements (the latter requiring double lead equalization) are given. Comprehensive "baseline" measurements were made on four well-trained subjects (three were pilots) using a specially built Controlled Element Computer and an on-line Describing Function Analyzer. The resulting data includes: tracking errors, describing functions (and derived loop closure and model-fitting parameters), remnant, critical instabilities, and reaction times. Remarkably simple correlations are shown between the critical instability and various other closed-loop dynamic performance metrics. These tests and results constitute the foundation for a series of continuing experiments on effects of environmental stresses, and workload.

PART II, PSYCHOPHYSIOLOGICAL CORRELATES

During a comprehensive set of tracking tasks for three orders of controlled element with four trained subjects (reported separately), a number of simultaneous psychophysiological measurements were made. These included: electrocardiogram, "instantaneous" heart-rate, breath flow, electromyograms, average grip pressure, and palmar skin resistance.

These data showed consistent increases in the neuromuscular tension indicators during tracking. Breathing was usually faster and shallower. The average heart rate increased for only two of the four subjects, but distinct increases in the cardiac "sinus arrhythmia" were noted, which were completely correlated with breath flow.

*This work was performed as part of Contract NAS2-4405 from the NASA-Ames Research Center, Man-Machine Integration Branch.

INTRODUCTION

To measure the dynamic response properties of the manual controller during long-duration space flights, a battery of tracking tasks and measurement is being developed. These measures are intended to be:

- Meaningfully related to the combined performance of the man-machine system, yet universally applicable and capable of extrapolation to tasks other than those tested.
- Sensitive to significant changes in the visual-motor system.
- Efficiently obtained* so as to require neither excessive on-board equipment, nor an undue amount of data transmission.

The tracking task battery developed herein is an outgrowth of a large body of manual control research performed previously (e.g., see Refs. 1 and 2 and the bibliographies therein) and, in particular, of the "critical instability" concepts developed in Refs. 3 through 6.

The underlying principle employed is to stretch one or more of the basic tracking skills to its measurable limit, thereby obtaining dynamic performance measures relatively free from confounding and variable stability and workload "safety margins."

This paper summarizes the results of our recent work in the following areas, which are prerequisites to operational use of the task battery:

a. Evolution and validation of a test battery which is both efficient and meaningful;

b. Development of testing devices and data reduction schemes (hardware and software) to perform the tests and interpret the results;

c. Development of baseline (normative) test data against which to test individual deviations due to the influence of various stress factors.

For convenience, the order of presentation will be b) and a) in Part I, then c) in Part II.

Complete details, rationale and background for this material are given in Ref. 1, which will be issued as a NASA Contractor Report. Only the key results are presented here, accompanied by the relevant conclusions and recommendations as they naturally occur.

---

*Efficient measurements are those requiring a minimum run length and/or fewest trials to establish stable estimates of the data parameters.

PART I — TEST DEVELOPMENT AND VALIDATION

A. MEASUREMENT TECHNIQUES AND APPARATUS

To measure human operator dynamic response properties in a scientifically useful way, more than mere error scores for a first- or second-order tracking task are required. An extensive and sophisticated technology of human operator dynamics now exists (Refs. 2 and 17 and bibliographies therein) and several types of efficient models* are available to structure such measurements and to coalesce the data. Furthermore, a number of time-domain or frequency-domain techniques for making efficient measurements of the desired properties are available. Based on considerable experience and consideration of the key dynamic response properties of interest (e.g., man-machine system stability, accuracy, bandwidth or response time, noise, and adaptability; see Ref. 1), we have chosen to measure the following:

- Tracking error performance (error variance relative to input variance).
- Various quasi-linear describing functions (and resulting parameters thereof) using a random-like forcing function.
- Remnant; on an overall basis (coherence) and power spectrum.
- Dynamic effective time delay via the Critical Instability Task.†
- Choice reaction time; using the same visual-motor paths.
- Various psychophysiological correlates, such as: cardiovascular functions, neuromuscular activity, and arousal measures (see Part II herein).

Some of the key dynamic response parameters measured during this program are given in Table I along with relevant "autopaced critical instability" measures. The latter are intended for use as clinical, minimal-processing

---

*Efficient models are those representing the major behavioral properties with the simplest forms.

†"Critical" and "subcritical" tracking tasks employ an unstable controlled element of the form: $Y_C = K_n\lambda/s^{n-1}(s-\lambda)$, where n is the task order. The operator stabilizes the system via a compensatory display of error. For the "Autopaced Critical Task" the degree of instability, λ (in rad/sec), is increased from a small value, rapidly at first, then gradually until the critical point of man-machine closed-loop instability is reached; this yields the critical instability score, λ_c. For various reasons, best λ_c results are obtained with no forcing function except the operator's own remnant noise. "Subcritical tracking" is performed with similar elements using a constant level of instability about one-third the critical value ($\lambda \doteq \lambda_c/3$). This forces concentrated operator attention but leaves ample stability margins to always complete the typical 100 sec runs.

TABLE I. SOME IMPORTANT HUMAN TRACKING PARAMETERS AND RELEVANT CRITICAL TASKS

PARAMETER	SYMBOL	TRACKING RELEVANCE	CRITICAL TASK APPLICATION
Effective delay times	τ_e	Indicator of effective control bandwidth: τ_{e_1} for proportional-control action τ_{e_2} for rate-control action Index of workload/stress capability Indicator of neuromuscular tonus Limits closed-loop stability Limits closed-loop bandwidth	 $Y_{c_1} = \lambda/(s-\lambda)$; $\lambda \rightarrow \lambda_{c_1}$ $Y_{c_2} = \lambda/s(s-\lambda)$; $\lambda \rightarrow \lambda_{c_2}$ $\lambda_c - \lambda_{c_o} = f$ (stress) $\omega_{c_{max}} \sim 1/\tau_e \sim \lambda_c$
Low frequency lead generation ability	T_L $K_R/\Delta\tau_{12}$	Determines integrity of lead channel, e.g., whether pure inertia-like vehicle can be controlled Indicator of rate-perception integrity	$\Delta\tau_{12} = \tau_{e_2} - \tau_{e_1} = f(T_L) \sim (1/\lambda_{c_1} - 1/\lambda_{c_2})$ $K_R/\Delta\tau_{12} \sim \lambda_{c_2}/K_{c_{2_{min}}}$
Operator remnant	Φ_{nn_e} , σ_{nn_e} $\rho^2_{a_e}$	Determines disturbance-free, or "at rest," variability of operator. Its minimum tracking errors attainable under "best conditions."	$\Phi_{nn_{min}} \sim \Phi_{ee}\Big\|_{Y_{c_1}} \rightarrow \sigma_{n_e}$; $i = 0$ $\overline{e^2} \doteq \sigma^2_{n_e} + \frac{1}{3}\left(\frac{\omega_c}{\omega_i}\right)^2$; $i \neq 0$

indices of the basic tracking properties they reflect. One purpose of the experiments was to validate these correlations.

A versatile but portable test apparatus, shown in Fig. 1, was developed to administer the tasks and measure, on-line, the describing function and remnant data (see Ref. 1 for details). The Controlled Element Computer (CEC) mechanizes any of 12 instantly selectable controlled elements: $\pm K_1$, $\pm K_1/s$, $\pm K_2/s^2$, $\pm K_3\lambda/(s \pm \lambda)$, $\pm K_4\lambda/s(s \pm \lambda)$, and $K_5\lambda/s^2(s-\lambda)$. It also provides for various "critical task" autopacing schemes, steady tracking provisions, remote and local control functions, and self-test provisions. The CRT Display is a compact unit with an error bar which moves up and down through a ±5 cm range for compensatory tracking, and displays the critical instability score to the subject after each autopaced trial. The Control Stick Assembly includes an isometric (force) stick held between thumb and fingers and pressed up or down by the wrist control muscles (carpi group). Grip pressure (normal to the control force) is measured on the stick as a muscle-tension index. Computer "operate" and "reset" controls are available for self-administered tests.

The Describing Function Analyzer (DFA), provides a random-like input from a sum of five sinusoids of adjustable amplitude, and computes on-line the finite Fourier transform (the real and imaginary vector components) of a selected signal in the loop (usually error) at the five input frequencies. It also measures the mean-squared and mean of the selected (error) signal, and the mean-squared value of another selectable signal (usually control force). The final describing functions, error variance, and relative remnant coherence are computed off-line via a handy time-sharing computer program. The details of the very accurate and efficient scheme for performing these measures are beyond the scope of this paper, and are described in Ref. 16.

The Describing Function Analyzer proved even more useful than originally hoped because its describing function, relative remnant, and performance data were accurate and rapidly available. The effort involved in obtaining the DFA-derived describing functions via time-shared computer processing was much less than the comparable off-line data processing (involving tape recordings, digitizing, digital time series analysis, and further subsequent manipulations).

A second set of this apparatus is now available for "piggyback" experiments on various other NASA programs (e.g., it is currently in use on a 90-Day Sealed Chamber Experiment).

B. SUBCRITICAL AND CRITICAL TRACKING TASK DEVELOPMENT

A set of first-, second-, and third-order subcritical tracking task parameters was developed to serve as a basic battery for future research. Figure 2 and Table II give the specific combinations of instability, display and control gains, and input characteristics. Based on the consistent results obtained with four experienced subjects (to be described later), we conclude that these test parameters provide a challenging and useful set of tasks which are reasonably easy to learn (a possible exception is the third-order task, which appears nearly impossible at first). The demands of each task are sufficiently high to yield near-asymptotic values for the operator's intrinsic tracking parameters, yet sufficient performance reserve is available to permit investigations of performance decrements resulting from various stress conditions.

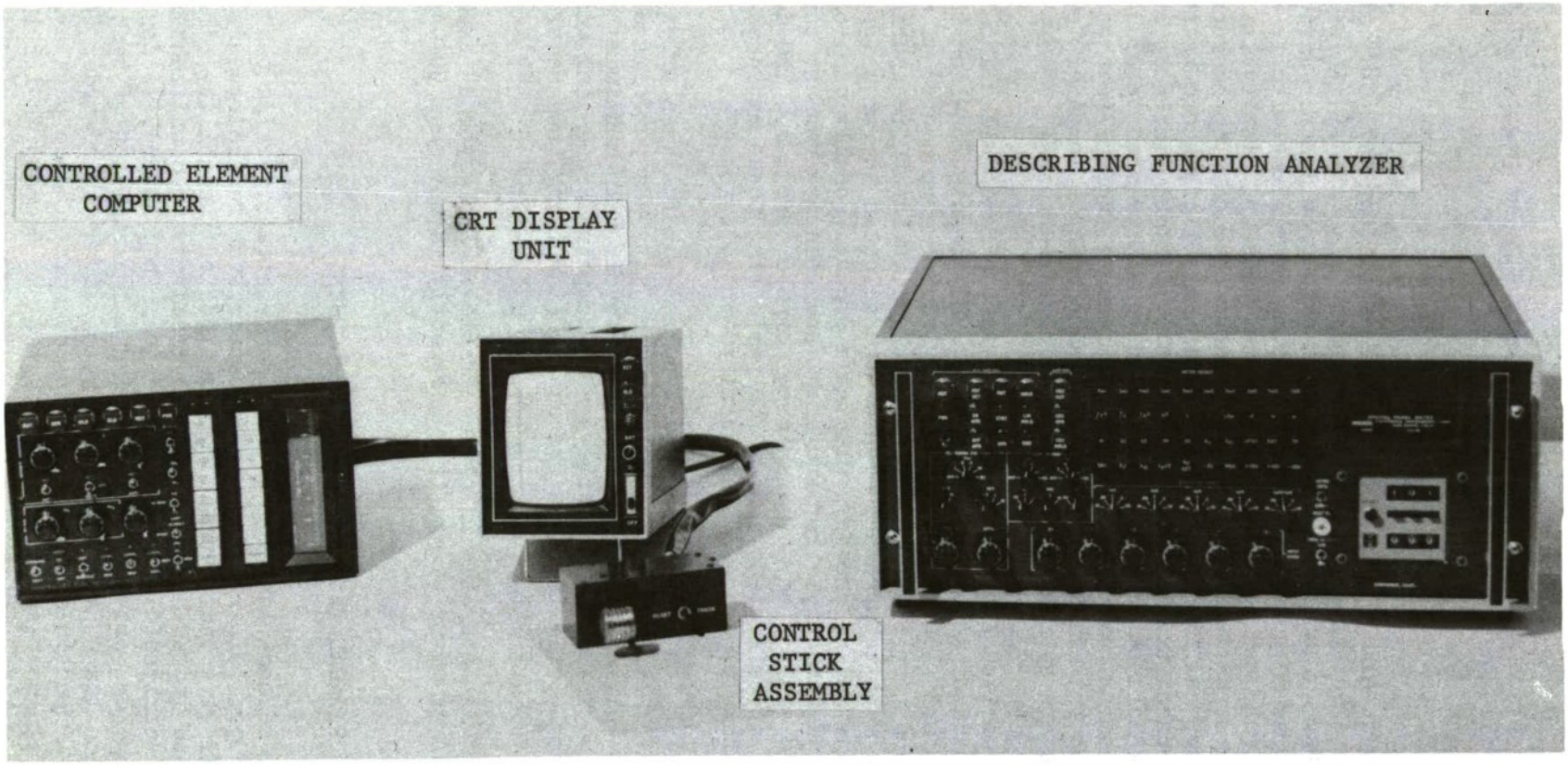
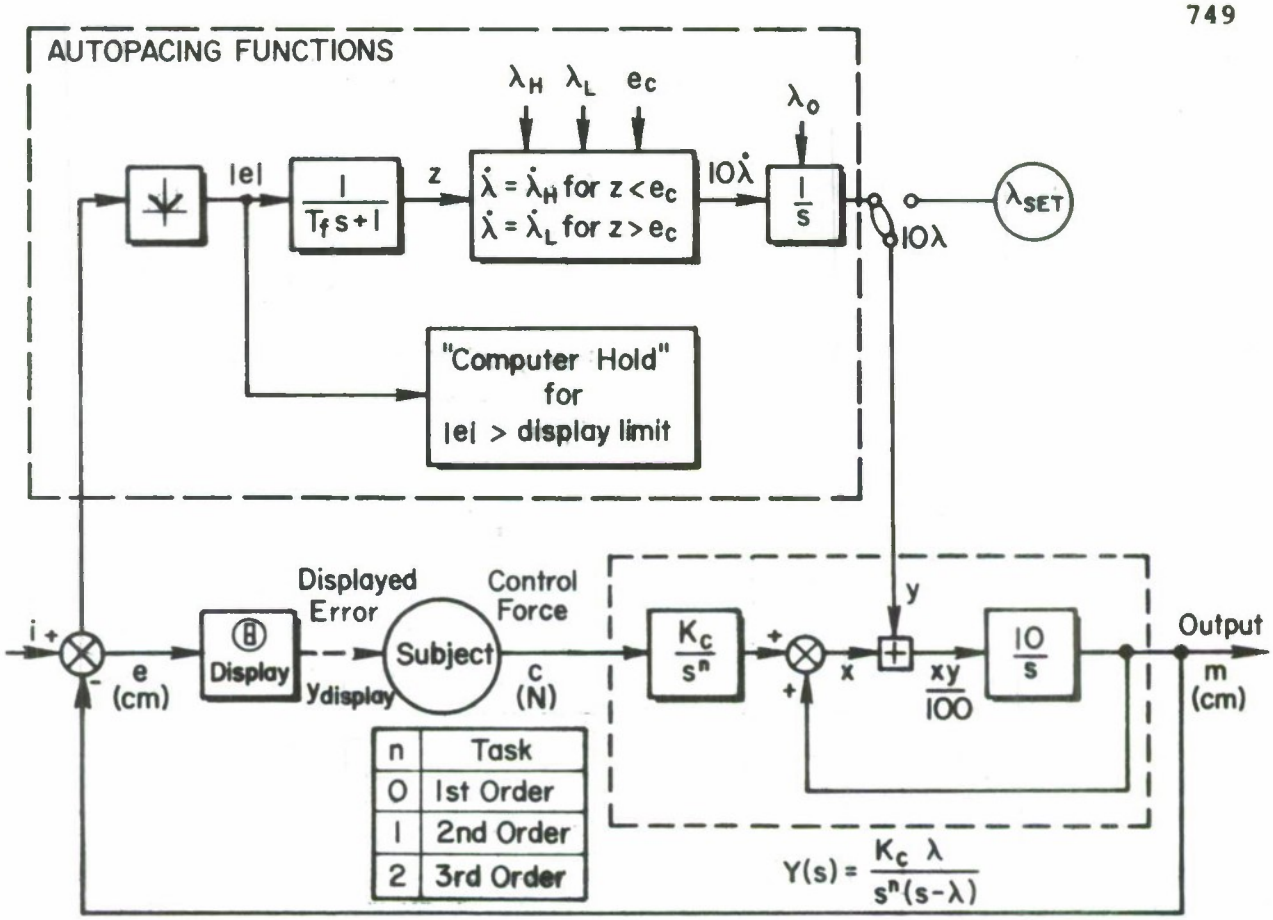


Figure 1. Psychomotor Test Battery Equipment



Autopacing Parameters	Units	1st Order	2nd Order	3rd Order
λ_0 - initial value of unstable root	rad/sec	1.0	.5	.25
$\dot{\lambda}_H$ - initial (high) λ rate	rad/sec^2	.20	.10	.02
$\dot{\lambda}_L$ - final (slow) λ rate of increase, $\dot{\lambda}_L = .25\ \dot{\lambda}_H$	rad/sec^2	.05	.025	.005
e_c - error criterion for switching from $\dot{\lambda}_H$ to $\dot{\lambda}_L$	cm	.5	1.0	1.0
T_f - error criterion smoothing time constant	sec	1.0	1.0	1.0

Subcritical Tracking Parameters	Units	1st Order	2nd Order	3rd Order
λ_{Set} - controlled element unstable root	rad/sec	2.0	1.25	.30
ω_i - input bandwidth with second order or greater roll off	rad/sec	2.0	1.25	.30
σ_i - input rms amplitude	cm	.50	.50	.50

Control Stick and Display Gains		1st Order	2nd Order	3rd Order
K_c Control/Display Sensitivity	Development tests with MSI stick	.9 cm/N	2.6 $\frac{cm/sec}{N}$	2.6 $\frac{cm/sec^2}{N}$
	Final tests with STI stick	.7 cm/N	1.4 $\frac{cm/sec}{N}$	3.5 $\frac{cm/sec^2}{N}$
K_D Display viewing gain for 50 cm nominal viewing distance		1.0 $\frac{\text{degree visual angle}}{\text{cm display deflection}}$		

Figure 2. Critical and Subcritical Tracking Task Parameters

TABLE II

FORCING FUNCTION SPECTRA

$$\text{Spectral Envelope: } \Phi(\omega)\Big|_{\omega_k} = \left\{ \frac{\frac{(j\omega)^2}{\omega_1^2} + \frac{2\zeta_1}{\omega_1}(j\omega) + 1}{\frac{(j\omega)^2}{\omega_2^2} + \frac{2\zeta_2}{\omega_2}(j\omega) + 1} \right\}^2$$

$$\zeta_1 = \zeta_2 = 0.707$$

CONTROLLED ELEMENT		$Y_{c1} = \frac{2}{s-2}$	$Y_{c2} = \frac{1.25}{s(s-1.25)}$	$Y_{c3} = \frac{0.3}{s^2(s-0.3)}$
Spectral Shape Designation		II-A	II-B	II-C
Numerator break, ω_1 rad/sec		9	5.6	2.1
Denominator break, ω_2 rad/sec		2	1.25	0.3
Frequency (rad/sec)	Cycles per 100 sec run	Sine Wave Amplitudes* (cm)		
$\omega_0 = 0.1884$	3	—	—	0.666
$\omega_1 = 0.502$	8	0.494	0.570	0.236
$\omega_2 = 1.256$	20	0.460	0.402	0.0436
$\omega_3 = 3.015$	48	0.204	0.103	0.0160
$\omega_4 = 6.282$	100	0.0543	0.0376	0.0143
$\omega_5 = 10.46$	166.67	0.0306	0.0300	—

*Amplitudes scaled to give mean square = 0.25 cm^2.

Very consistent and accurate describing function data were obtained on the first- and second-order tasks. Reasonably accurate describing functions were obtained while tracking with a third-order subcritical task (which requires double-lead equalization to be stable). However, time traces show that third-order equalization is obtained by an extremely pulsive control action (Fig. 3c), resulting in an error signal which is mostly remnant ($\rho_E^2 \leq 0.2$). In view of the possible requirement for this type of control equalization (e.g., during manual reentry backup control with stability augmenters failed) we recommend further investigation of the pulsed control models needed to more accurately represent and describe the second-order lead behavior.

Building on the experience with the already developed (Ref. 3) autopacing scheme for obtaining the first-order critical instability (λ_{c1}), similar autopacing parameters were developed for the second- and third-order critical instabilities (λ_{c2}, λ_{c3}). Table II describes the appropriate combinations of

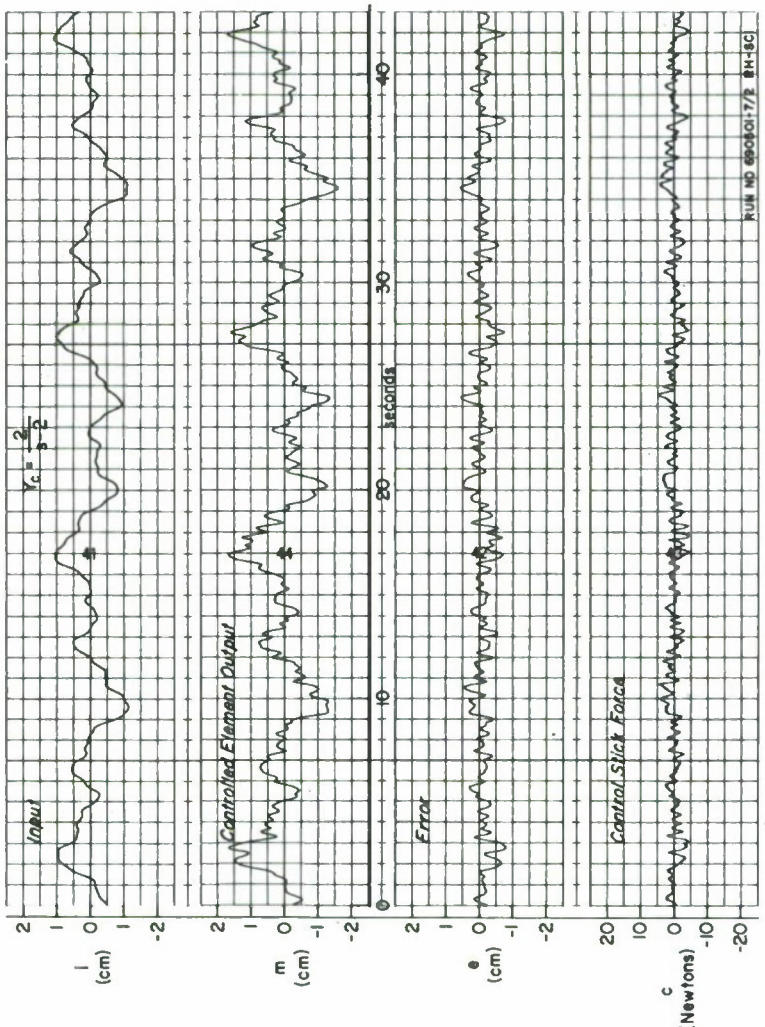


Figure 3. Time Traces of Tracking Signals

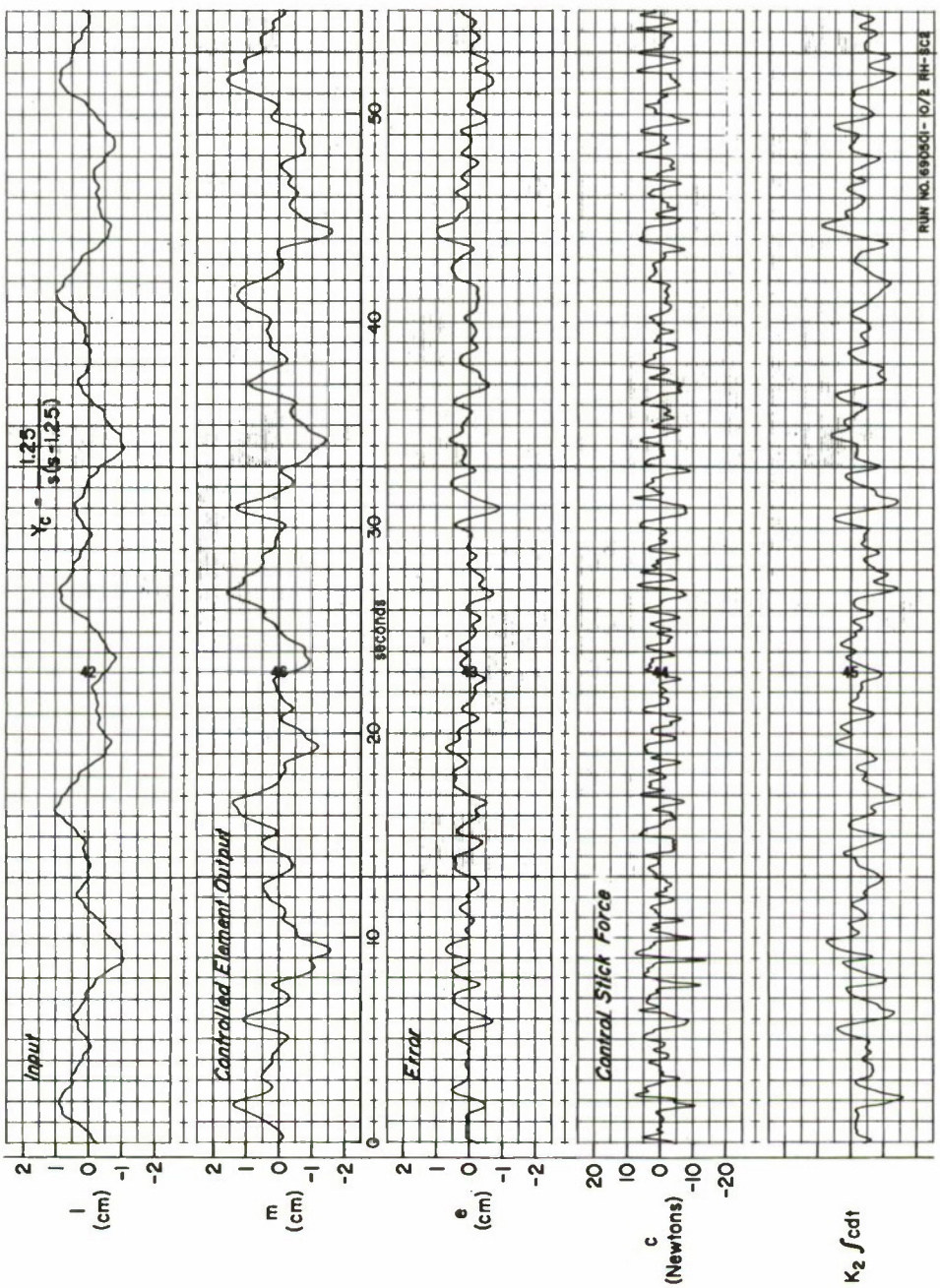


Figure 3 (Continued)

Figure 3 (Concluded)

initial and final rates of instability increase, rate-switching error criteria, and initial instability. These values were shown to yield reasonable run lengths and reliable critical instability scores (see Ref. 1). A general principle was found for selecting the final rate-of-increase of instability: "When approaching its unstable level, λ should not be increased more than a few percent per period of the dominant closed-loop oscillation."

Based on successful application of these autopaced tasks to the four subjects in this program, as well as in training other members of the staff, it is concluded that development is complete for the first-, second-, and third-order critical instability tasks (which demand zero-, first-, and second-order operator lead equalization). For comparable results in various labs, the critical task parameters should be within ±25 percent of those given in Table II.

C. TEST VALIDATION

1. Training

Four skilled subjects (three instrument-rated pilots with different varying commercial and military flight experience, and one nonpilot technician highly trained on tracking tasks; ages from 24 to 43 years) were carefully trained and measured by a comprehensive battery of autopaced and subcritical tracking tasks and step reaction times. In addition, a number of psychophysiological measurements were made (see Part II). The trends in the training data on performance and critical instability were roughly similar among the four subjects, and the selected test parameters seemed to be well suited to all of them. Stable asymptotes were reached in first- and second-order tracking, but for two subjects learning was still taking place on the very difficult third-order task. This experience is consistent with past results, showing that controlled elements requiring much equalization take longer to learn than those requiring minimal equalization. Typical asymptotic levels of critical instability and standard deviations were as follows:

Order	N trials	Range of $\bar{\lambda}_c$	Typical $\bar{\lambda}_c$	Std. Dev., σ_λ	$\sigma_\lambda/\bar{\lambda}_c$, %
First	5	5.0 - 6.5	6.2	0.3	5%
Second	5	2.7 - 3.9	3.5	0.3	8%
Third	5	0.5 - 1.2	1.0	0.13	15%

2. Basic Loop Closures

The measured open-loop describing function and injected remnant data were similar, in general characteristics, for all four pilots, although each differed in the levels of specific parameters. Each pilot's data for a given set of three successive runs was extremely consistent, as shown in Fig. 4. There were small but self-consistent differences among the pilots in gain and phase crossover frequencies (Figs. 5 and 6), while a few larger differences were observed in phase margin, crossover frequency, and error performance (Fig. 7a). The relative remnant, as measured by the error coherence (Fig. 7b), and normalized remnant spectrum were remarkably consistent among subjects for a given order task.

A comparison was made between the data for these four subjects and the data taken three years ago for a single subject over a wider range of subcritical and critical task variables (Refs. 3 and 4). The tie-in shown in Fig. 6 between past and previous data is excellent in every respect, thereby validating the earlier work for one subject and giving some confidence in the previously explored effects of various control stick types, the effects of λ on the adapted parameters, and effects of input size and spectral shape on the adapted describing function and remnant. From the agreement among the present subjects and the good tie-in with the past research, we conclude that all subjects adopt the same basic strategy, similar results can be obtained in different facilities using different mechanizations of the same test, and investigations of other task variables may be made using a small number of well-trained subjects.

3. Model Fits

The describing functions for the four subjects were reasonably well fit by the simple "extended crossover model" (α-model of Ref. 1) in the region near the crossover frequency (see solid line in Fig. 4). The fitted pilot model parameters, K_{OL}, τ_e, α, and T_L, were generally similar for most operators, and the differences were consistent with corresponding changes in performance (Fig. 8). The α-model parameters were also consistent with the values from prior experiments.

Systematic deviations from the simple α-model fit were found for all subjects in the present experiment, and are consistent with those found in previous experiments: a high frequency rise, known to be due to high-order neuromuscular modes whose phase contribution is accounted for by τ_e in the α-model; and a low frequency rise, best fitted by a lag-lead equalization in the pilot describing function. In addition to the previous explanation (in Ref. 10) of this low frequency lag-lead as a neuromuscular phenomenon, it can be shown that the same effect may be caused by conscious compensation (trimming out of residual drift errors) by the pilot. This is called the "parallel pseudo-integrator model," and is explained in Ref. 1. Replacing the α term by a low frequency lag-lead to account for this pilot behavior (or the use of the "precision model" of Ref. 10) produces satisfactory analytical estimates of the error spectrum and total error. From analyzing these data we conclude that models and parameters more refined than the simple α-model may be required if one is investigating tracking performance differences among skilled operators. The fact that a low frequency lag-lead term is required to fit the data from three separate experiments with subcritical-type controlled elements indicates that more definitive research of these phenomena is certainly warranted.

Achieving stability in second-order subcritical tracking requires predominantly lead equalization. The present results show that the operators did not use the maximum lead known to be achievable, but rather some intermediate value which would achieve better performance without undue sacrifice of stability margins. Two of the subjects showed low or even negative values of α for the second-order case, because the "best" T_L overcorrects the mid-frequency phase.

The increase in τ_e from 0.12 to 0.20 sec from first- to second-order subcritical tracking is consistent with past results in Refs. 4 and 10. For third-order subcritical tracking a further penalty was evident ($\tau_e = 0.4$-0.6 sec),

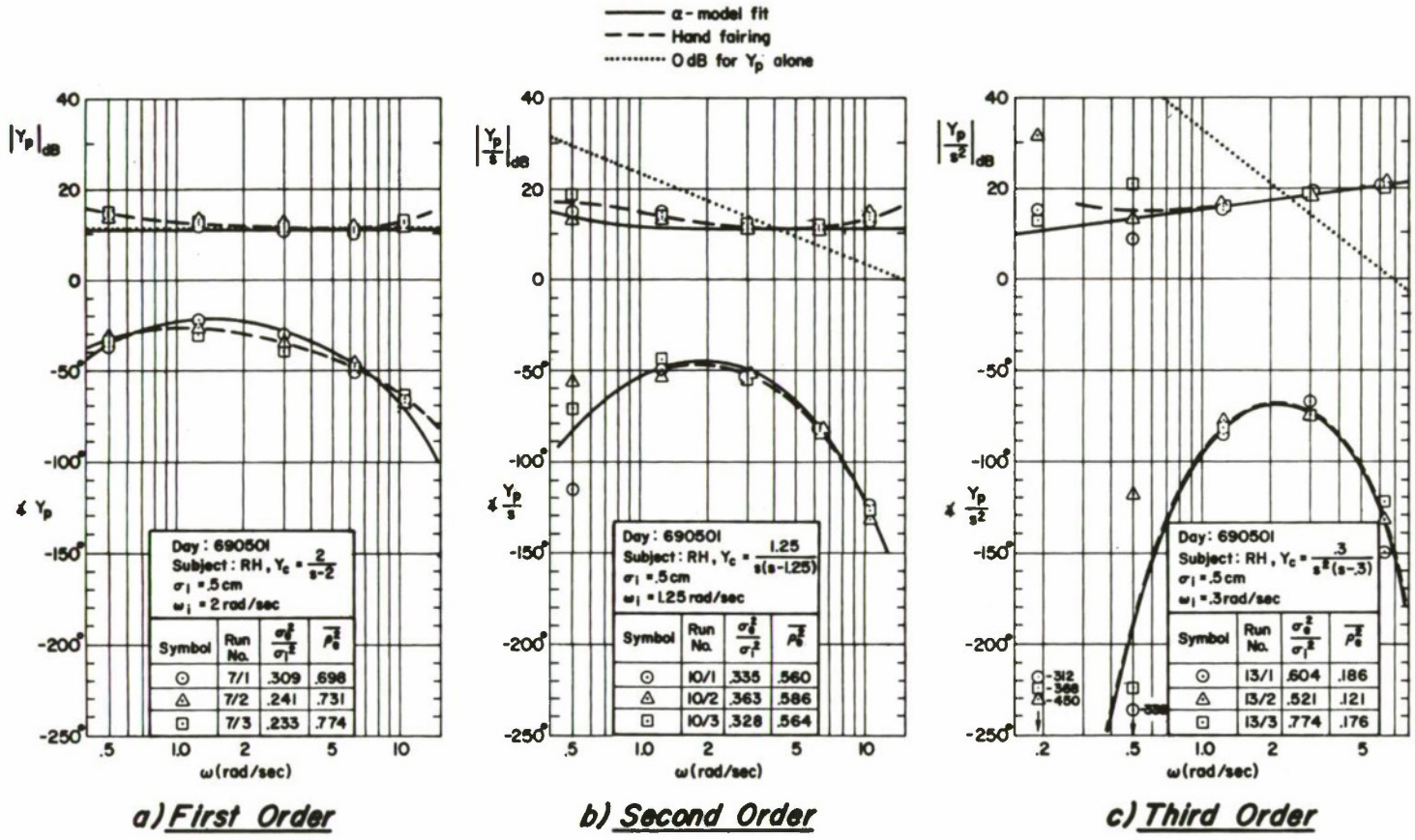


Figure 4. Typical Pilot Describing Function Data, Hand Fairings and Model Fits

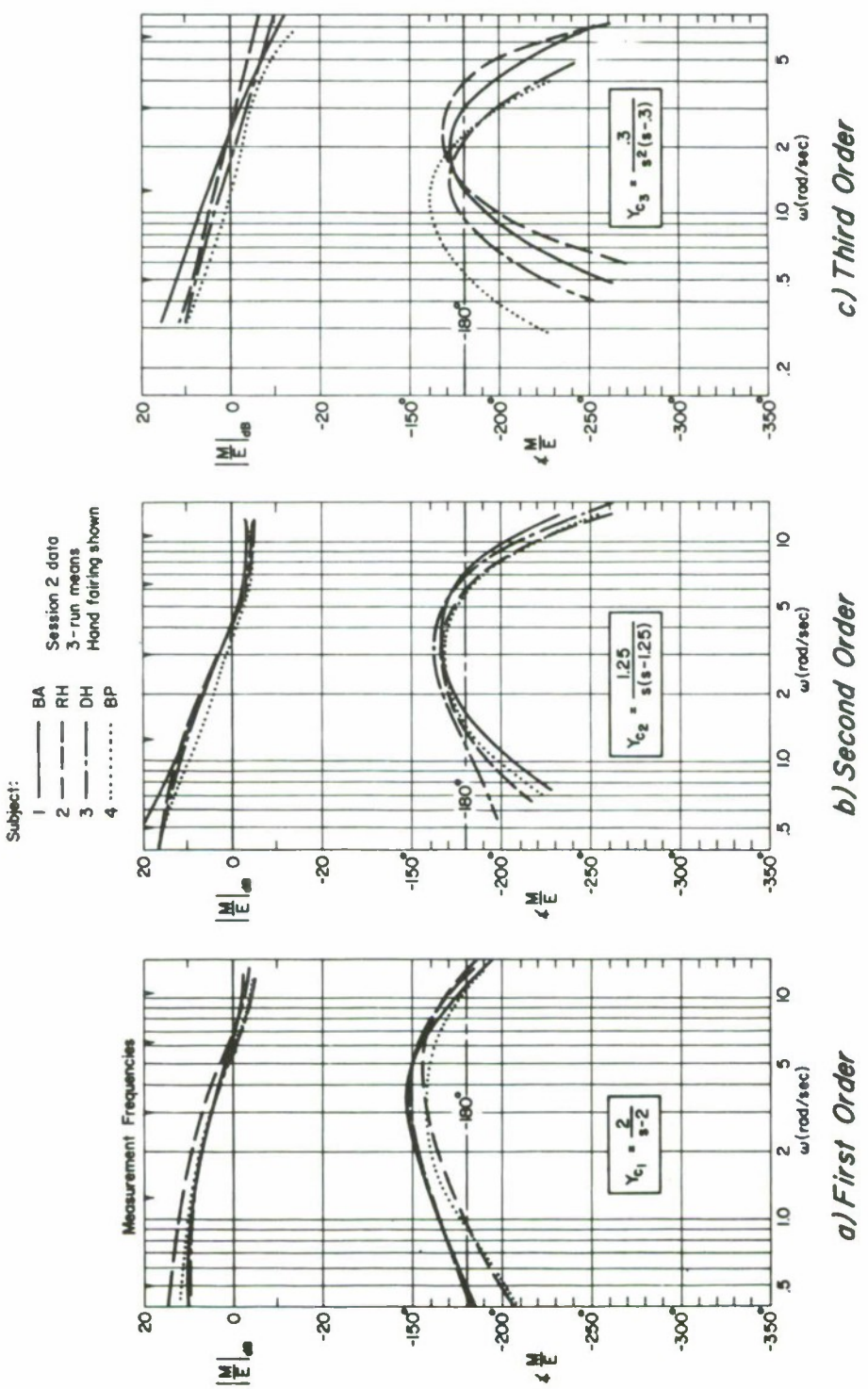


Figure 5. Comparison of Faired Open-Loop Describing Functions for 4 Subjects and 3 Subcritical Controlled Elements

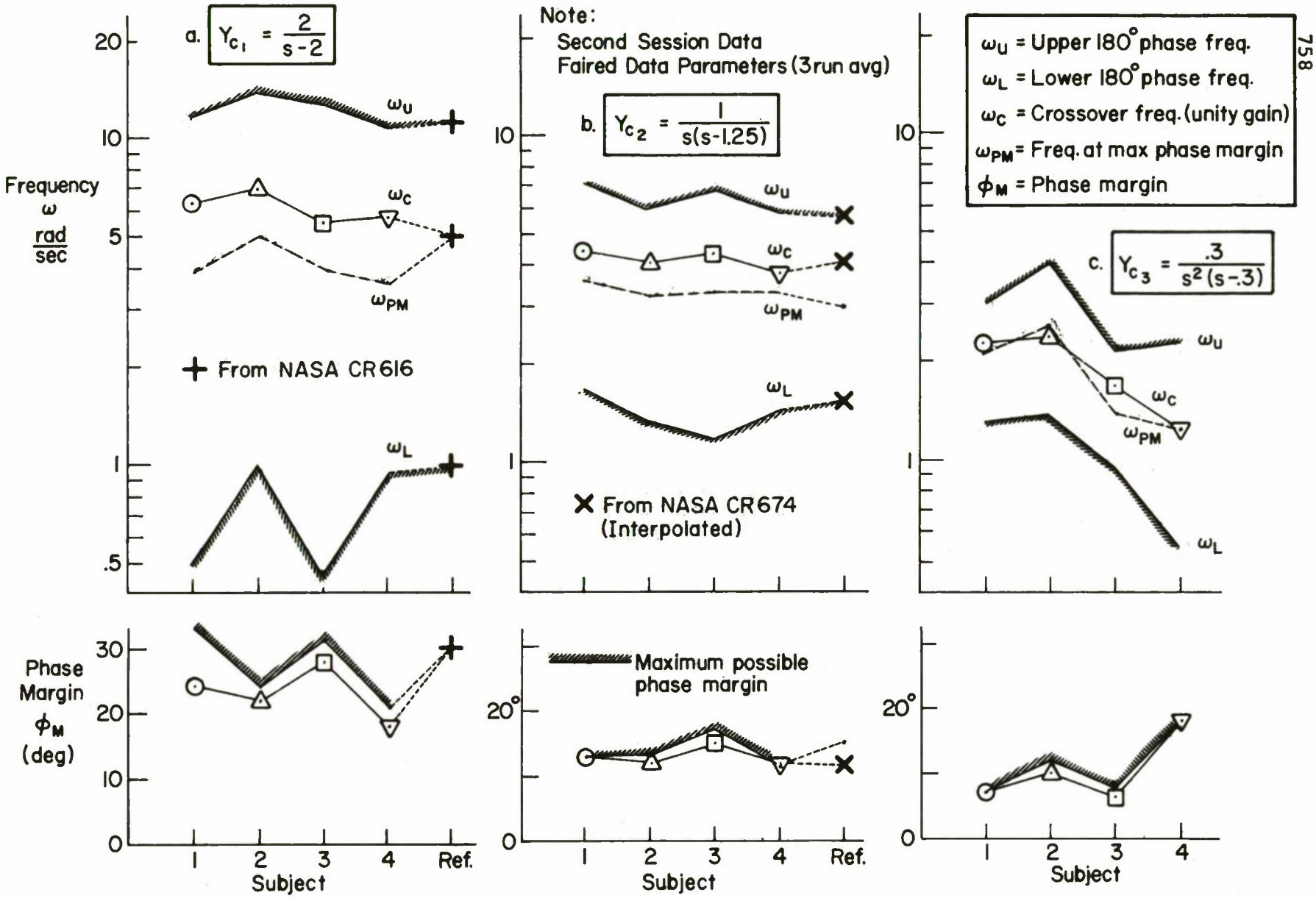


Figure 6. Faired Describing Function Data Parameters

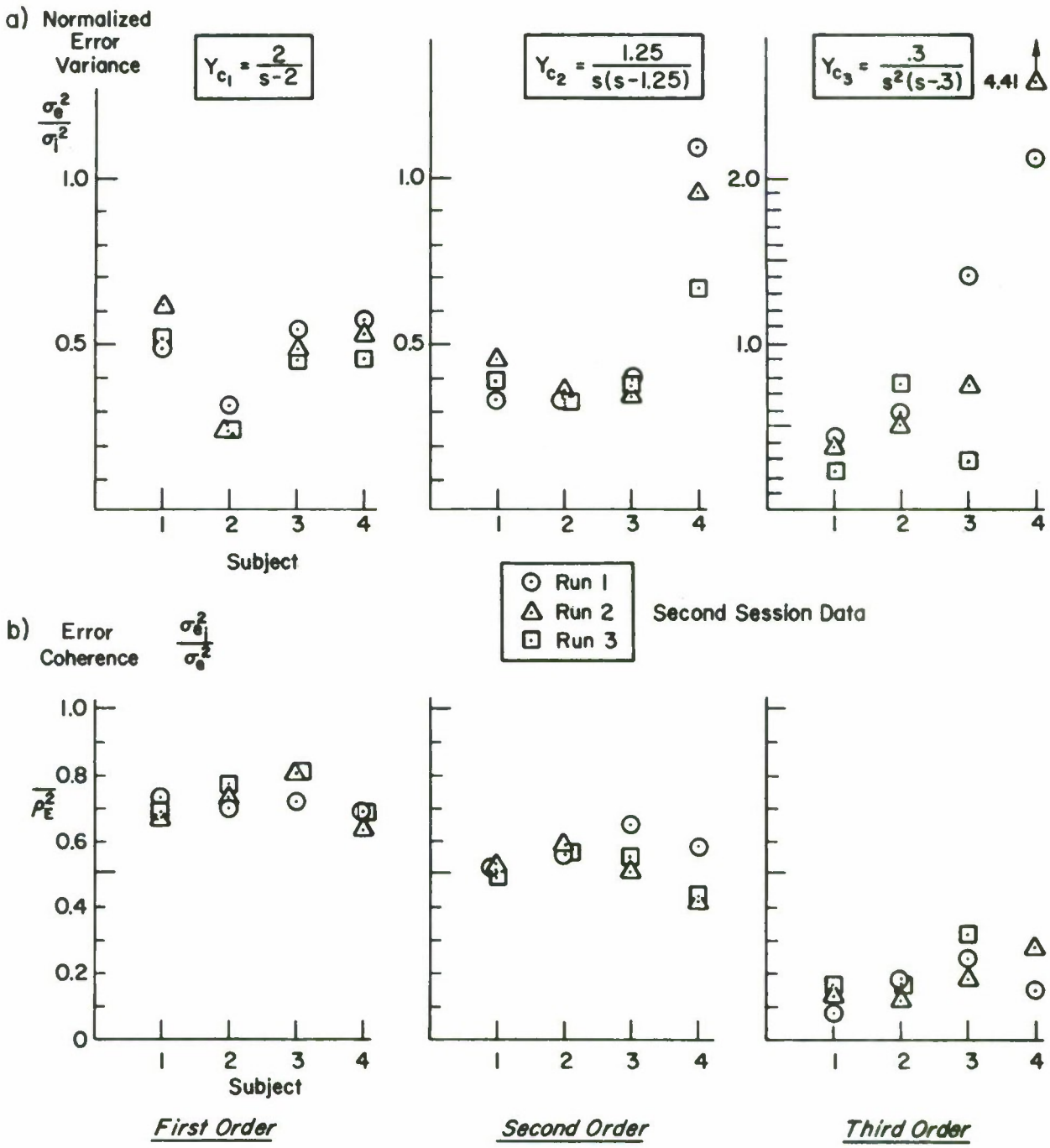


Figure 7. Comparison of Mean Square Tracking Error and Overall Coherence

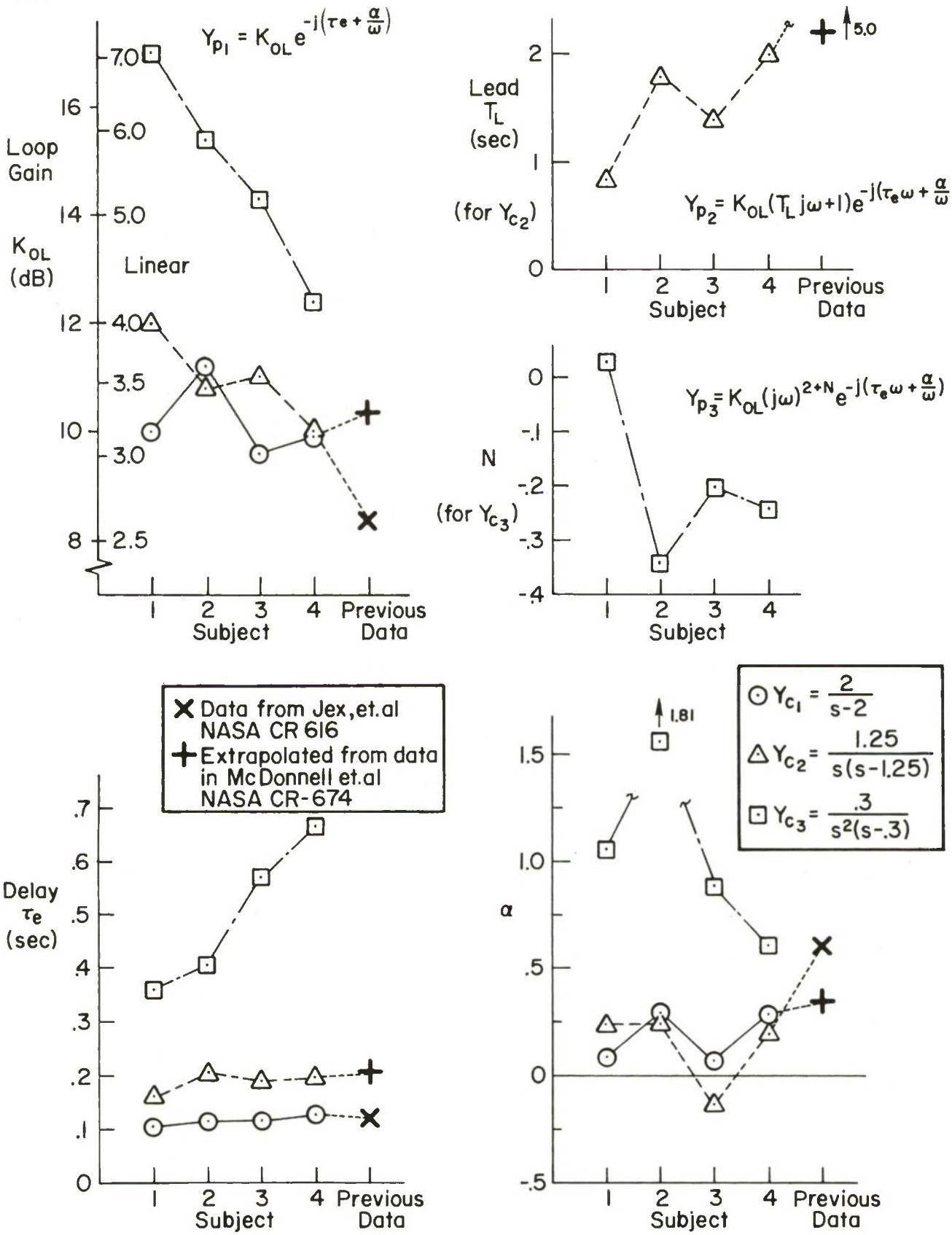


Figure 8. Fitted Human Operator Model Parameters

implying that a similar internal processing scheme may be at the root of both rate and acceleration equalization. These trends validate the concept that the time delay penalty for generating lead is a more appropriate measure of lead-generating ability than the lead time constant itself.

An interesting relationship was noted between the individual levels of gain adopted (measured by the gain crossover frequency, ω_c) and the upper unstable gain (measured by the phase crossover frequency, ω_U). As plotted in Fig. 9, the data for a given order element tend to coalesce along constant ratios of ω_U/ω_c of 2.0 for first-order and 1.6 for second- and third-order subcritical tracking. Because of the unity slope of $|Y_{OL}(j\omega)|$ near crossover, these ratios correspond to stability (gain) margins of 6 dB and 4 dB, respectively. These small and consistent margins suggest that the crossover region tracking behavior may be governed primarily by stability considerations (and the ensuing implications on error overshoot), rather than tracking error minimization.

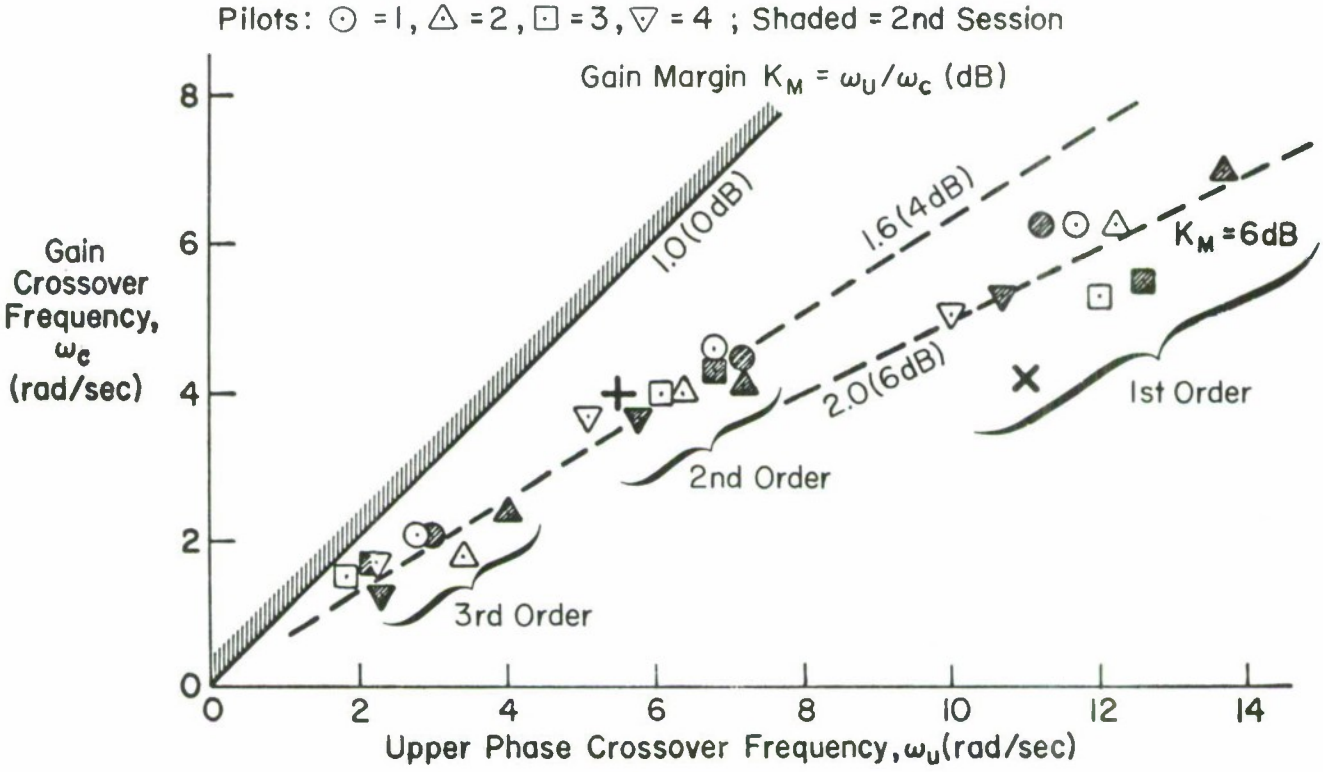


Figure 9. Correlation between Gain and Phase Crossover Frequencies

4. Remnant

Examination of the error and control signal power spectral densities showed that the input-correlated power was readily distinguishable from the level of the remnant (with the third-order cases being marginal). Even in the third-order case, where the operator was known to be pulsing irregularly, the control remnant spectra were quite smooth. When transformed to an equivalent injected remnant input at the error point and normalized by the mean-square error (i.e.,

per Refs. 7 and 8), the resulting plots of $\Phi_{nn'_e}$ for each of the four operators coalesced well for a given order of controlled element (Fig. 10). Although no detailed investigation was made of these remnant data, they do seem consistent with the earlier subcritical tracking experiments in Refs. 4 and 10 and with the normalized remnant data of Refs. 7 and 8.

An interesting finding is that, at any given controlled element order, the overall error coherence, ρ_E^2, was remarkably consistent among each of the four subjects, as shown in Fig. 7b, despite significant differences in their describing function parameters and overall performance levels ($\rho_E^2 \doteq 0.7$, 0.5, and 0.2, respectively, for first-, second-, and third-order tasks). This constancy of ρ_E^2 could be because the remnant is mostly multiplicative processing noise, but previous research on subcritical tracking with different inputs (Ref. 4) has shown that the remnant does not vanish as $i \rightarrow 0$. Because of the importance of operator remnant under small input conditions, we recommend further detailed research in this area.

All in all, the steady tracking results showed that quite accurate and consistent describing function data could be efficiently obtained on-line, with simple and portable apparatus. A sound base of normative data has been built up to serve as a reference for future experiments.

D. CORRELATIONS BETWEEN λ_c SCORES AND TRACKING BEHAVIOR

A number of correlations were made between measures of stability during steady tracking (such as effective delay time, phase and gain crossover frequencies) and the independently measured critical instability scores. It has been shown (Ref. 4) that the inverse delay time, τ_e^{-1}, is about 1.3 times the critical instability, λ_c. As shown in Fig. 11a, the correlations of τ_e with λ_c are excellent for variations among first-order cases, fair among second-order cases, and marginal among third-order cases. More importantly, the same correlation line fits all three controlled elements, and it is consistent with the earlier preliminary correlation of Ref. 4. A least-squares fit to all these data is given by the simple expression:

$$\tau_e^{-1} \doteq 1.1 + 1.2\lambda_c \quad ; \quad R = 0.98 \tag{1}$$

From the obviously good correlation overall and the fairly good correlation between individual variations, we conclude that λ_c is a reasonably good estimator of the limiting value of τ_e for a given type of operator equalization. From the similarity in correlation among the three controlled elements we also conclude that decrements in λ_c for controlled elements requiring equalization, compared with the first-order score (no equalization), are good indicators of the dynamic response delay penalties paid for various degrees of lead equalization. Good correlation was also obtained between λ_c and ω_U (the upper unstable frequency). Since ω_U sets an upper limit to the bandwidth of the man-machine system, the tie-in between λ_c and system bandwidth has thus been validated. As shown in Fig. 11b, a common correlation line fitted the data for all three controlled elements, giving as the best fit:

$$\omega_U \doteq 1.1 + 1.72\lambda_c \quad ; \quad R = 0.92 \tag{2}$$

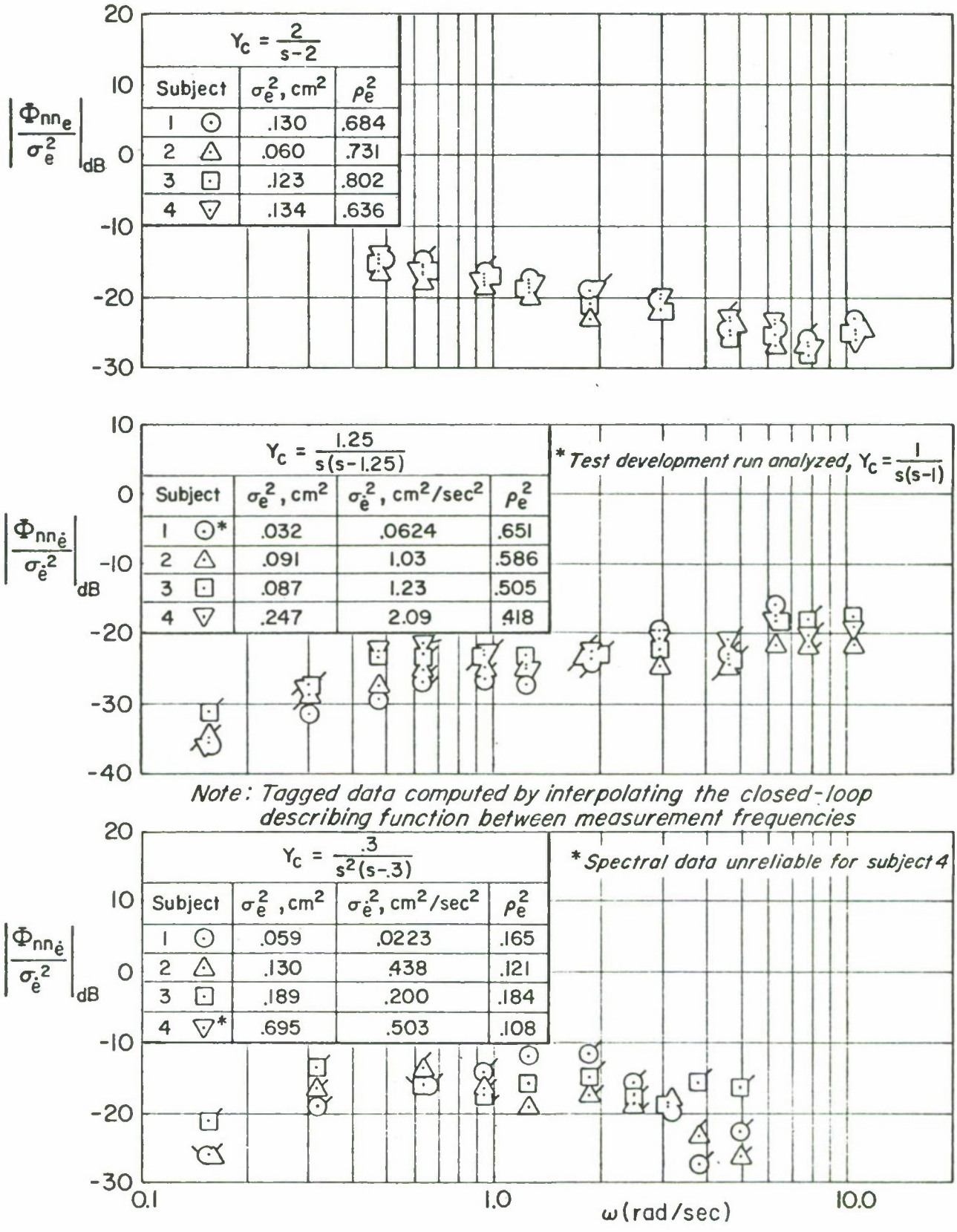
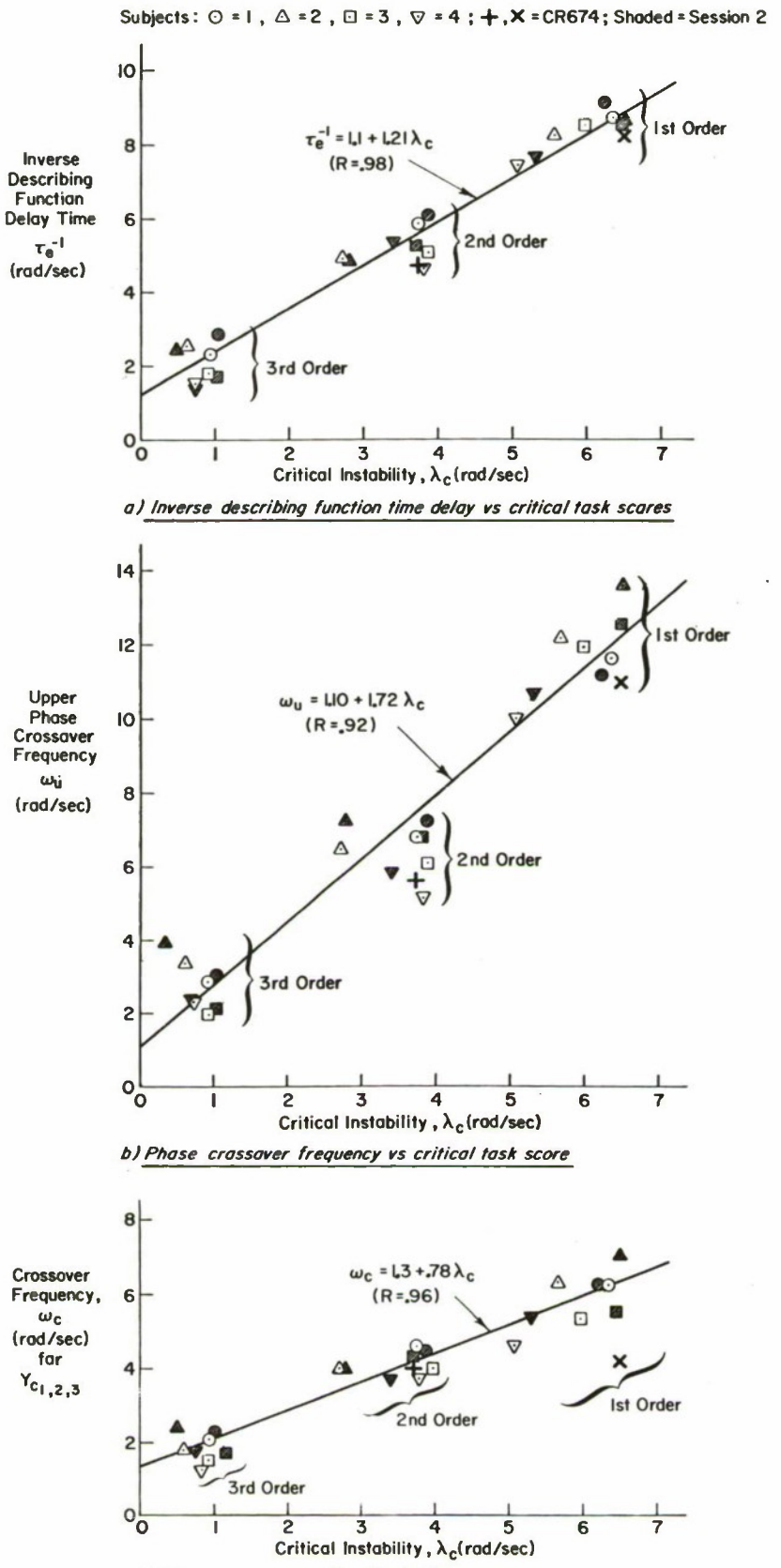


Figure 10. Equivalent Injected Remnant Spectra Referred to Error Signal



a) Inverse describing function time delay vs critical task scores

b) Phase crossover frequency vs critical task score

c) Gain crossover vs critical instability scores

Figure 11. Correlations Between Tracking Parameters and Autopaced Critical Instability Scores

Because the crossover gain, and thus crossover frequency, are limited by the gain margin to some fraction of ω_U (see Fig. 9), it was expected that ω_c should show some correlation with λ_c also. However, other factors, such as performance optimization and remnant suppression, might cause more variability in ω_c than ω_U. The ω_c data from the present experiment show a surprisingly good correlation with λ_c, as shown in Fig. 11c. The approximate relationship found for subcritical tracking is:

$$\omega_c \doteq 1.3 + 0.78\lambda_c \quad ; \quad R = 0.96 \tag{3}$$

The correlation between λ_c and performance measures, such as $\overline{e^2}/\overline{i^2}$, is not nearly so good, due to the considerations mentioned above and the low frequency lag-lead equalization, as mentioned previously. If all of the closed-loop performance factors were to be revealed by individual critical-type tests, then separate tests would be needed to define both the low frequency lag-lead equalization and the neuromuscular system peak as effectively as the present test defines τ_e. Since this approach would add several tests and parameters, we feel that it is easier to simply measure the five-frequency describing function whose five vector descriptors define "exact" fits to the data.

Comparisons of λ_c with the step reaction time (the appropriate statistic is RT^{-1}) showed poor correlation (Fig. 12). (Correlations between RT^{-1} and τ_e^{-1} were similarly unimpressive.) Part of the difficulty is due to the intrinsically higher variance of the disjunctive reaction time measurement, and part is due to the fact that different visual-motor pathways may be involved. The critical task instability score shows a ratio of $\sigma_{\lambda_c}/\overline{\lambda_c}$ of roughly 5 percent for N = 5, while the disjunctive reaction time shows a ratio of $\sigma_{RT^{-1}}/\overline{RT^{-1}}$ of approximately 19 percent for N = 10. The current results validate the discussion of reaction data contained in Ref. 4. Furthermore, the disjunctive reaction time data alone cannot give a decrement related to the time delay penalty for generating higher order lead equalization. We conclude that, even though step reaction time measurements may be fairly easy to make, their higher intrinsic variability and uncertain relevance to dynamic response behavior while tracking render RT a poor test (estimator) for τ_e.

Taken across the board, these measurements on three skilled pilots and one nonpilot, plus the good tie-in with prior exploratory data, show that the critical instability scores are excellent indicators of the closed-loop man-machine system bandwidth and of the effective time delay while tracking. We now feel that the foundations have been well enough established to use critical instability scores for a wide number of clinical applications, without concurrent measurements of describing function parameters. For example, decrements in λ_c, resulting from substitution of a different form of display or control for the ideal CRT and force stick units of the critical task can be used as reliable indicators of the effective time delay penalties incurred by the nonideal devices, and the achieved level of λ_c is a good index of the limiting bandwidth of the system. Finally, the rapid learning and excellent stability of the λ_c scores permit very efficient experimental designs, in terms of minimizing the number of subjects, training time, and testing time.

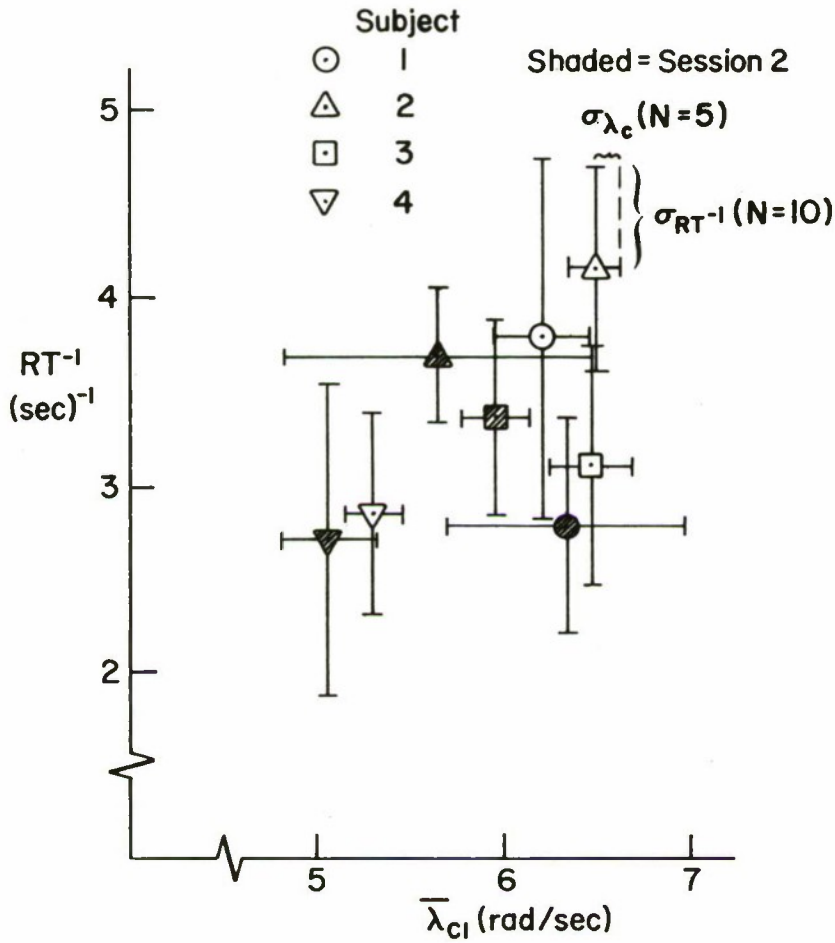


Figure 12. Correlation Between Inverse Reaction Time and First-Order Critical Task Scores

PART II — PSYCHOPHYSIOLOGICAL CORRELATES

A. APPARATUS

During the two final data sessions of the reference experiment described in Part I, a number of psychophysiological measurements were made in order to establish the baseline levels under ideal situations and to determine which measurements would be most useful in future experiments. Since no universally accepted package (or even a functional specification) for the measurements was available, we selected and adapted several existing techniques for use in this program:

GP — The stick grip pressure (thumb-to-forefinger force transverse to the direction of the control action) is measured by a transducer in the stick, and the mean value over the run is measured.

ECG — A pair of electrodes on the sternum (breastbone) is used for electrocardiogram measurement, along with a very high-impedance preamplifier.

HR — The instantaneous heart rate is measured by commercial cardiotachometer using the ECG R-wave leading edge as the trigger.

BF — An indication of the breath flow in and out of one nostril is obtained by a simple nasal thermistor unit. An equalizing circuit was developed to compensate for the thermistor lag. Breathing frequency was computed from this signal.

EMG — Rectified and filtered electromyograph activity, picked up by bipolar electrodes, was recorded at each of three sites: 1) and 2) a pair on the active arm on the dominant agonist-antagonist muscles (the carpi-radialis longus and carpi-ulnaris muscles which move the wrist); 3) at a corresponding site on the passive arm. High-impedance preamplifiers were used in conjunction with specially developed and matched rectifier filters to process these signals. In some cases a weighted sum and difference of the active limb EMG signals was used (corresponding to the average tension and force commands, respectively), while in other cases the average activity during a run was used.

PSR — Palmar skin resistance was measured between electrodes on the back and palm of the passive hand. Special circuitry was developed to insure constant current operation.

Ground — An under-tongue signal ground was used for all signals. This procedure eliminated much environmental electrical noise, compared with an ear-lobe ground, and insured purely nasal breathing.

The setup of the transducers and tracking station is shown in Fig. 13. Very clean signals were obtained throughout the test sessions, as shown in Fig. 14. Along with other tracking variables, the psychophysiological data were recorded on 14-channel, 1-7/8 ips FM magnetic tape for permanent filing. (See Ref. 1 for details of these measures; transcriptions of selected runs can be arranged by writing to the authors.) These data constitute a unique and comprehensive set of simultaneous tracking data and psychophysiological measurements on several skilled-pilot subjects, with three orders of controlled element.

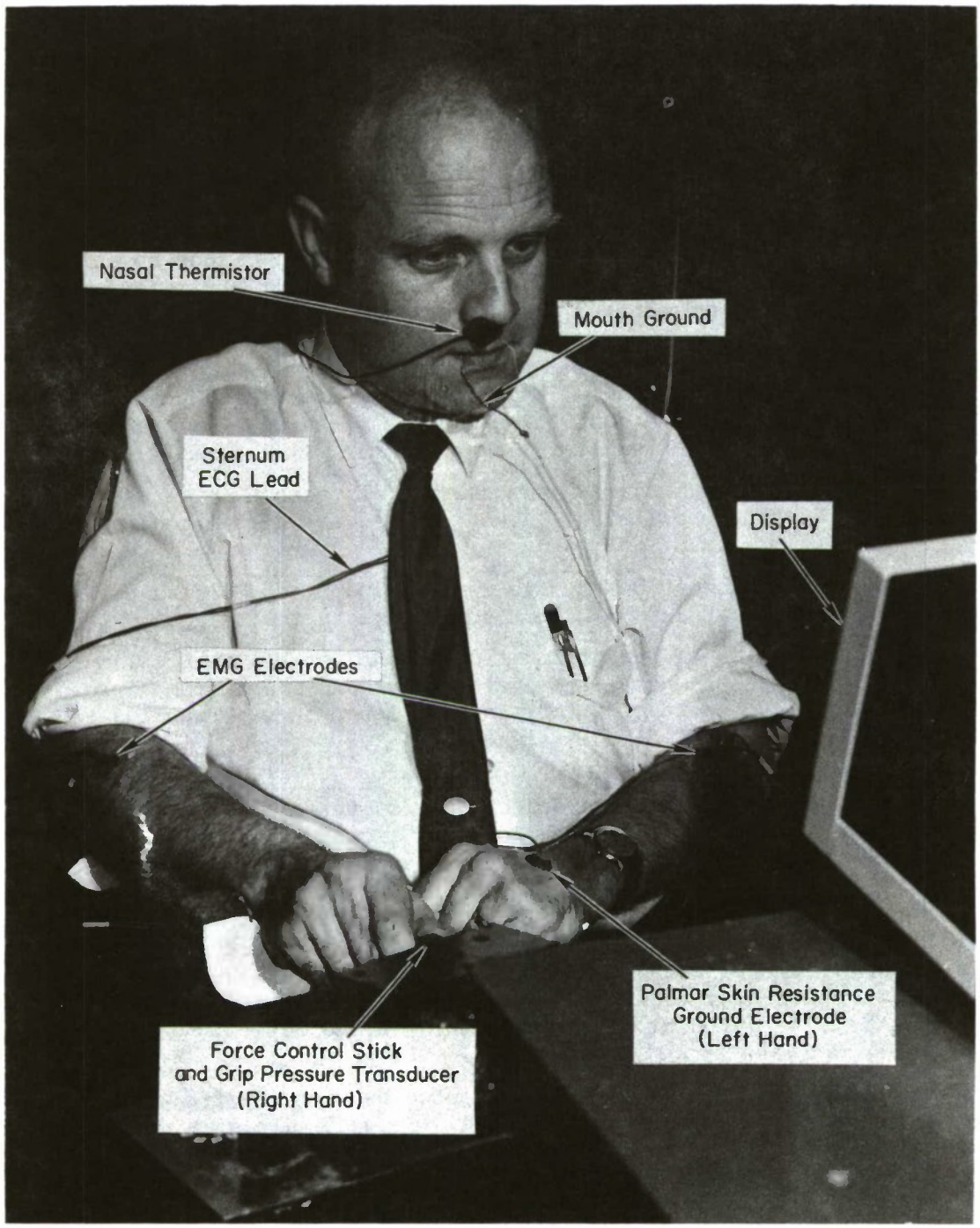
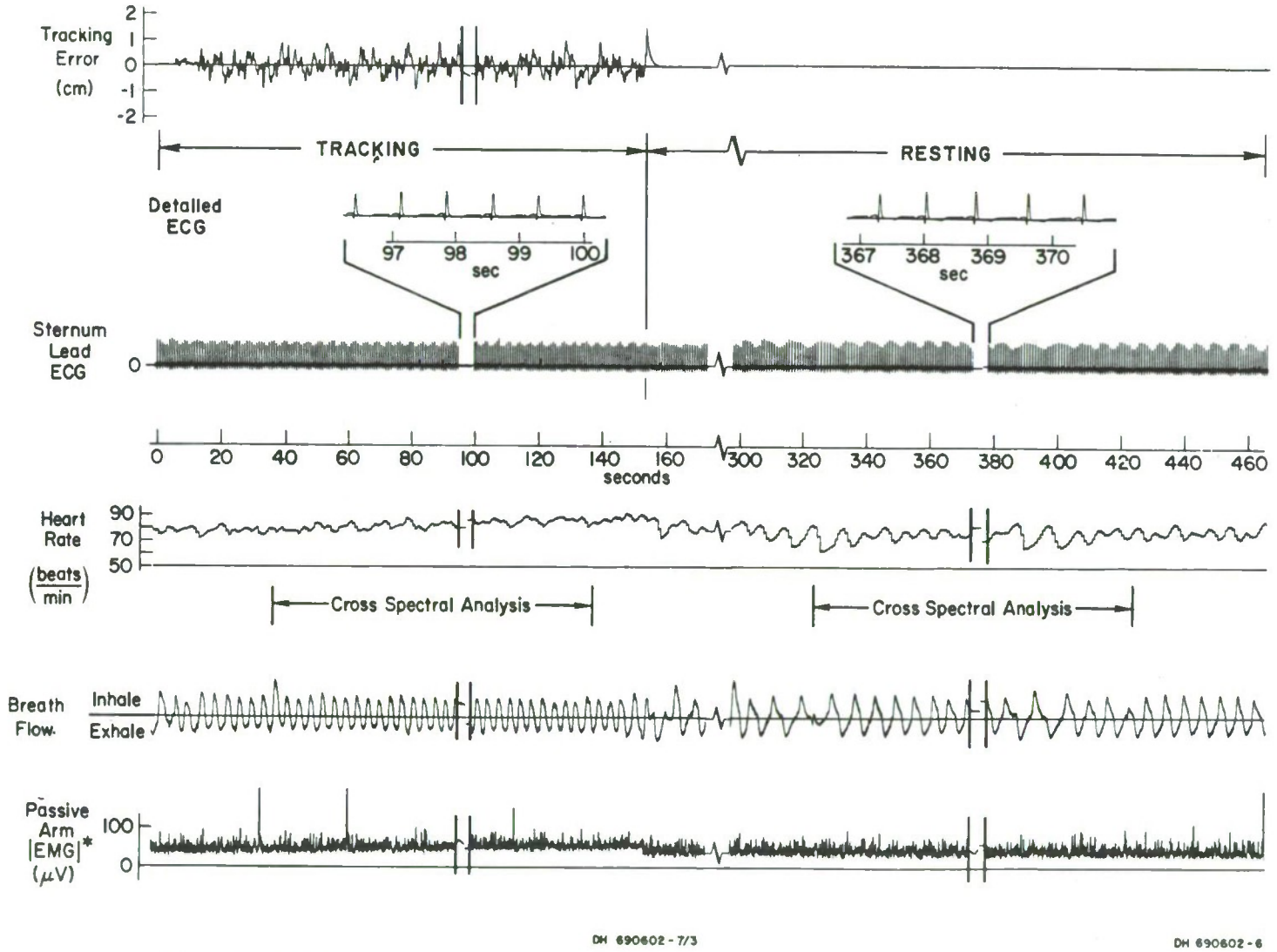


Figure 13. Psychophysiological Instrumentation



*Note : Raw EMG's were rectified and smoothed with a second order filter ($f_n = 5$ Hz, $\zeta = .7$).

Figure 14. Typical Time Histories of Psychophysiological Data

B. CORRELATIONS

1. Neuromuscular Correlations

All four subjects showed an increase in the average grip pressure (GP), active EMG, and passive EMG measures during tracking compared with resting (Figs. 14 and 15). Grip pressure traces showed a qualitative correlation with the passive EMG which suggests that GP can be used to measure systemic muscle activity. However, the mean GP is biased by the effects of the high frequency gripping pulsations which accompany the second- and third-order control tasks, as shown in Fig. 16. Further analysis between grip pressure and the various EMG signals should be performed to separate out the desired (low frequency) grip pressure trends from the (high frequency) artifacts of the control activity. We conclude that the average grip pressure (as measured on the Mark II Control Stick) is a useful indication of increased neuromuscular tension during tracking and hand-to-stick coupling, but that additional signal processing will have to be devised to prevent contamination by pulsive control gripping actions.

2. Cardiovascular Correlates

For two of our four subjects, increases in average heart rate did accompany tracking, as noted by other researchers (Refs. 11-13). Somewhat surprisingly, as illustrated in Fig. 17, the third-order task induced less heart rate increase than first- or second-order tasks, despite its subjectively higher difficulty. Because the average control force required for the third-order tracking was lower than the first- or second-order cases, our data suggest that the heart rate increase merely reflects the higher work output required of the former as compared with the latter. Two of the subjects (Subjects 1 and 2) did not show a significant change in heart rate between resting and tracking. Since the other subjects, as well as past research, generally show a rise in HR, these two anomalous cases warrant further investigation. The two nonsensitive subjects happen to be the best trained of the group, and one had previous experience with the somewhat awesome electrode array. Further investigation is clearly necessary before one can generalize about increases in cardiovascular functions during tracking tasks.

All subjects showed faster and shallower breathing during tracking (Fig. 14). As shown in Fig. 17, the average breathing frequency went from 13 to 20 breaths/minute. The qualitative effects of the breath flow signal could be accounted for by a rough limit on breath acceleration.

It is known that quasi-periodic "sinus arrhythmia" heart rate fluctuations are produced by breathing, and Clynes (Ref. 14) has developed a fairly complete analog computer model of this effect. Our data for Subject 3 show that during tracking the cardiac arrhythmia fluctuations decreased in amplitude, but increased in frequency, similar to the breathing changes. The rms fluctuation during a pair of typical runs decreased from 5 to 3.5 beats/min between resting and tracking, while the frequency of these fluctuations went from 8 to 17 cycles/min. To further analyze these effects, a cross spectral analysis was made between the equalized breath flow signal and the instantaneous heart rate

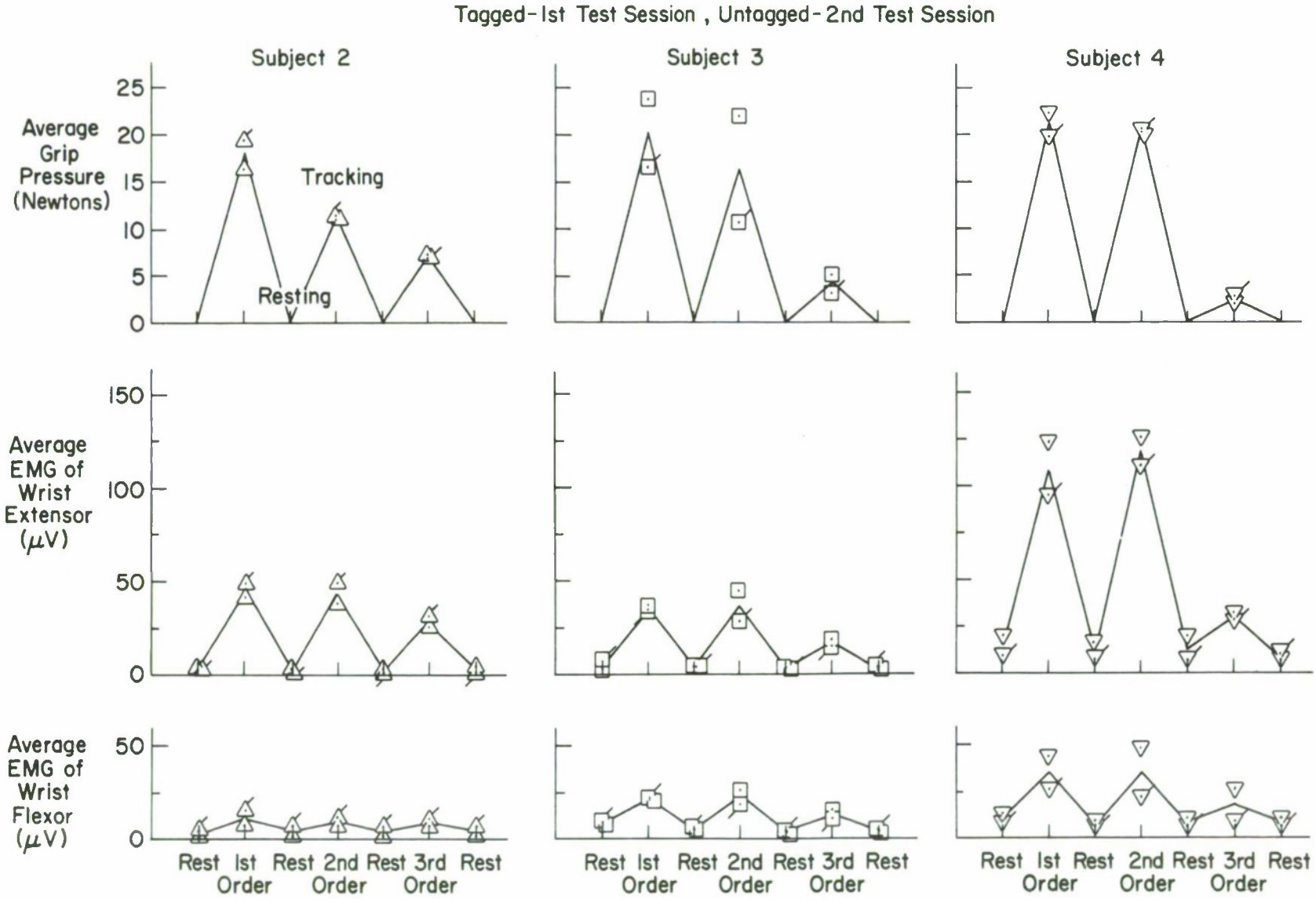


Figure 15. Average Grip Pressure and Active Arm EMG Levels During Tracking and Resting

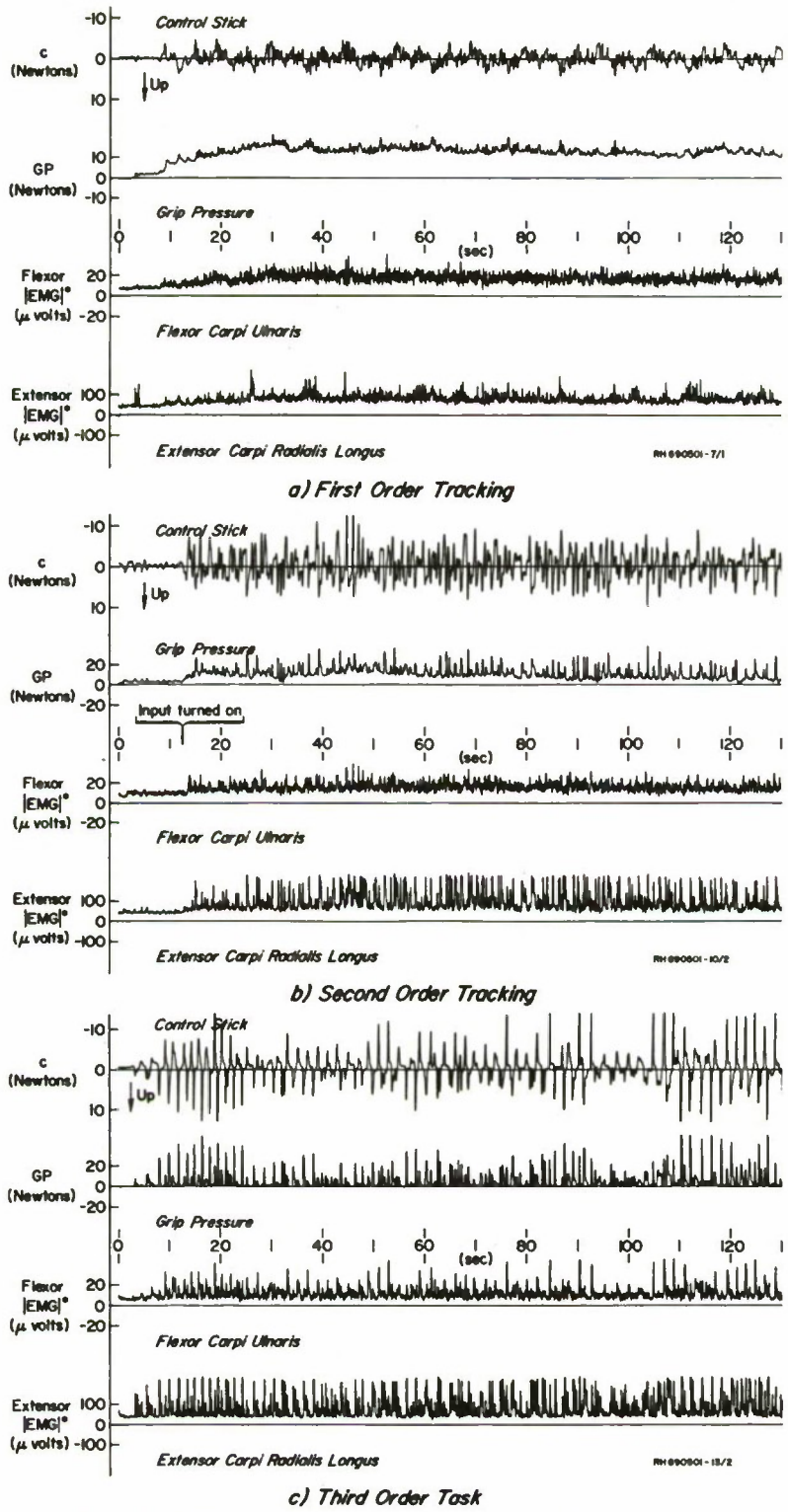


Figure 16. Typical Time Histories of Control Stick and Grip Pressure Forces, and Correlated EMG Activity

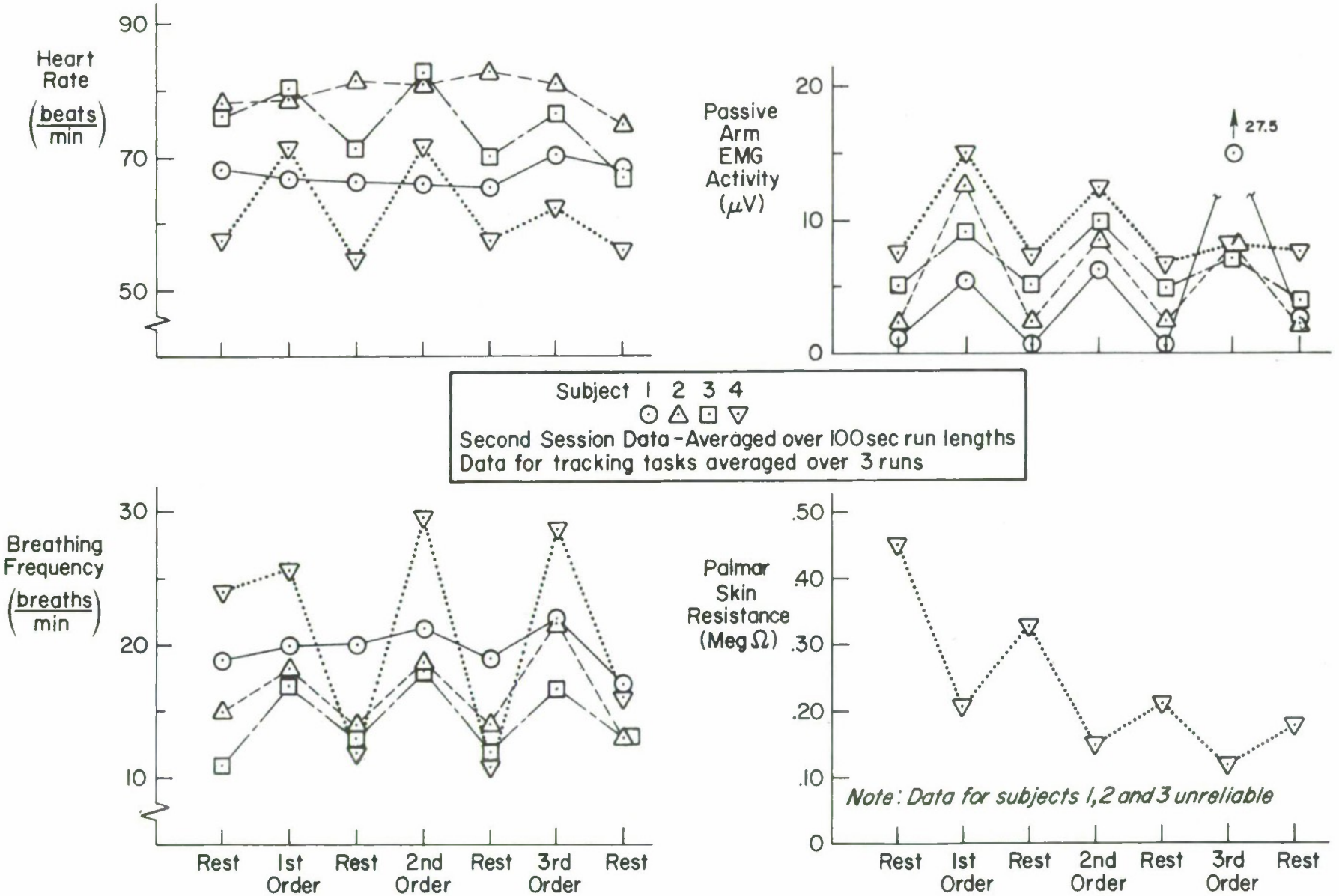


Figure 17. Psychophysiological Measurements During Rest Periods and Tracking Runs

signal (see Fig. 18). The power spectra of both BF and HR show peaks corresponding to the breathing frequency during both resting and tracking runs. The heart rate peaks are strongly correlated with the breathing frequency (coherency = 0.8 resting, 0.95 tracking) and the second harmonic of the breathing frequency is also seen in the heart rate spectrum. The portions of the heart rate spectra that are not correlated with breathing frequency remain fairly similar between resting and tracking conditions. We conclude that the majority of the sinus arrhythmia effect is related to the breathing interactions with the heartbeat, which is in accord with the Clynes model of this effect (Ref. 14).

Kalsbeek, in Ref. 15, has found that heart rate fluctuations decreased under high psychomotor loading conditions, and has proposed this phenomenon as a workload parameter. The present data on the heart rate fluctuations agree with his observations, and also suggest that the observed decrease in cardiac arrhythmia amplitude results primarily from faster, shallower breathing while tracking. Qualitative measurements of breath flow are easier to obtain than instantaneous heart rate measurements (e.g., compare BF obtained by a nasal thermistor unit or chest expansion potentiometer versus an ECG electrode and amplifier system and cardiotachometer). It would be desirable to see if Kalsbeek's extensive data (on cardiac arrhythmia decrements from various types of psychomotor loading) can be duplicated using simple breath flow measurements.

In referring back to Fig. 14 we observe that a breath flow "modulation" can be seen in the height of the ECG R-wave as recorded by sternum bipolar electrodes. It was not established whether this represents a true increase in the R-wave activation or was an artifact of the varying electrode-to-heart distance during breathing. Nevertheless, this observation does make it feasible to determine the breathing frequency for many past situations where breath flow was not measured, simply by counting the modulation peaks in the recorded sternum ECG signal.

3. Correlates of Arousal

Measurement of palmar skin resistance, PSR, was included because it is a commonly used measure of arousal or attentional interest. Because of some equipment difficulties, reliable PSR data were obtained on only one subject. The PSR did consistently decrease during tracking runs as compared with resting values, as shown in Fig. 17. This is consistent with the earlier results of Benson, et al., Ref. 11. We also noticed small, sudden decreases in PSR when the error signal started getting out of hand for one reason or another. No detailed investigation of PSR effects was made here because the data were intended mainly to provide a baseline against which to check the effects of future, more stressful situations.

4. Concluding Remarks

The general conclusion to be drawn from correlating several psychophysiological measurements with subcritical tracking is that appreciable variations occur in most which are fairly consistent with previous exploratory research in this area. None of the psychophysiological data indicated that the second- or third-order critical tracking tasks were more "stressful" than the first-

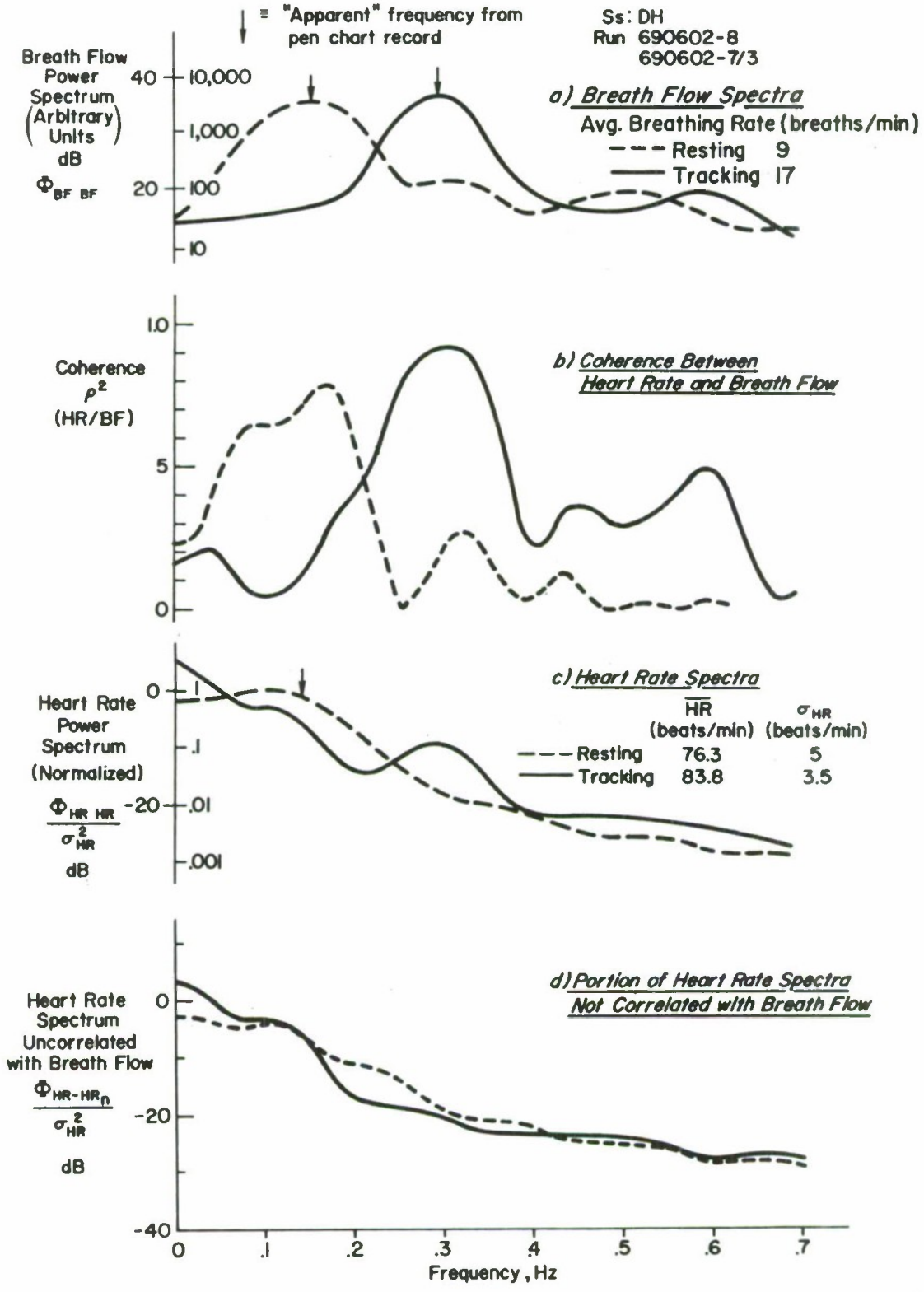


Figure 18. Breath Flow-Heart Rate Cross-Spectral Analysis Data

order tracking, yet the subjects generally agreed that the higher order tasks were intrinsically more "demanding."

In view of the difficulties that we had in assembling and tuning our psychophysiological measurement complex, we strongly recommend that a standard battery of such tasks be agreed on by workers in the manual control field. Some standardization is essential before data from various laboratories can be compared. The significant interactions observed between some of the measurements (such as between tracking and breath flow) must be understood and modeled before such measurements can be used as valid indicators of psychophysiological stress. Subcritical and critical tracking tasks of various orders and levels of instability are suggested as potentially valuable tools in this research because the operator's dynamic response behavior can be systematically controlled by the experimenter using the data and techniques we have presented here.

REFERENCES

1. Jex, Henry R., and R. Wade Allen, A Psychomotor Task Battery for Manual Control Performance: Development, Validation, and Some Psychophysiological Correlates, Systems Technology, Inc., Tech. Rept. 175-1, Jan. 1970 (forthcoming NASA CR-).

2. McRuer, Duane, and David H. Weir, "Theory of Manual Vehicular Control," Ergonomics, Vol. 12, No. 4, July 1969, pp. 599-633.

3. Jex, H. R., J. D. McDonnell, and A. V. Phatak, A "Critical" Tracking Task for Man-Machine Research Related to the Operator's Effective Delay Time. Part I: Theory and Experiments with a First-Order Divergent Controlled Element, NASA CR-616, Nov. 1966.

4. McDonnell, J. D., and H. R. Jex, A "Critical" Tracking Task of Man-Machine Research Related to the Operator's Effective Delay Time. Part II. Experimental Effects of System Input Spectra, Control Stick Stiffness, and Controlled Element Order, NASA CR-674, Jan. 1967.

5. Jex, H. R., J. D. McDonnell, and A. V. Phatak, "A 'Critical' Tracking Task for Manual Control Research," IEEE Trans., Vol. HFE-7, No. 4, Dec. 1966, pp. 138-145.

6. Jex, H. R., "Two Applications of the Critical Instability Task to Secondary Work Load Research," IEEE Trans., Vol. HFE-8, No. 4, Dec. 1967, pp. 279-282.

7. Levison, W. H., D. L. Kleinman, and S. Baron, "A Model for Human Controller Remnant," IEEE Trans., Vol. MMS-10, No. 4, Dec. 1969, pp. 101-108.

8. Jex, Henry R., and Raymond E. Magdaleno, "Corroborative Data on Normalization of Human Operator," IEEE Trans., Vol. MMS-10, No. 4, Dec. 1969, pp. 137-140.

9. 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., Tech. Rept. 170-2, July 1969 (forthcoming NASA CR-1569).

10. 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.

11. Benson, Alan J., Jo H. F. Huddleston, and John M. Rolfe, "A Psychophysiological Study of Compensatory Tracking on a Digital Display," J. Human Factors Society, Vol. 7, No. 5, Oct. 1965, pp. 457-472.

12. Roman, James, "Long-Range Program to Develop Medical Monitoring in Flight — The Flight Research Program — I," Aero. Med., Vol. 36, No. 6, June 1965, pp. 514-518.

13. Iguchi, Masakazu, "Manual Control Systems Including Two or More Operators," Bulletin of the Japan Soc. of Mech. Eng., Vol. 6, No. 24, Nov. 1963, pp. 696-703.

14. Clynes, Manfred, "Respiratory Control of Heart Rate: Laws Derived from Analog Computer Simulation," IRE Trans., Vol. ME-7, No. 1, Jan. 1960, pp. 2-14.

15. Kalsbeek, J. W. H., and J. H. Ettema, "Sinus Arrhythmia and the Measurement of Mental Load," Communication at the London Conf. of the Brit. Psychol. Soc., Dec. 1965.

16. Allen, R. Wade, and Henry R. Jex, A Simple Fourier Analysis Technique for Measuring the Dynamic Response of Manual Control Systems, Systems Technology, Inc., Paper 92, Mar. 1970.

17. Young, L. R., "On Adaptive Manual Control," Ergonomics, Vol. 12, No. 4, 1969, pp. 635-675.

IDENTIFICATION OF PILOT DYNAMICS WITH AND WITHOUT MOTION CUES

By
E. P. Salmon
J. T. Gallagher

NORTHROP CORPORATION
Aircraft Division
3901 West Broadway
Hawthorne, California 90250

SUMMARY

An investigation into the effect of motion cues on pilot describing functions has been initiated at the Northrop Aircraft Division. Results of an experiment performed during this investigation are reported in this paper.

The Northrop Large Amplitude 3-Axis Flight Simulator was used to provide a control task for a test pilot. The task was a single loop control closure measured as lateral compensatory tracking, (bank angle control). Five data runs were made using both the 5-degree-of-freedom motion and no motion capabilities of the simulator. A different set of aircraft dynamics was used for each data run. Pilot dynamics were assumed to be of the form $K\frac{\tau s + 1}{s + b}$, and values of K and τ were identified during the data runs. Rates of aileron and elevator deflection and task error were also measured. Although these measurements represent only a first attempt, they indicate that associated with fixed base simulation are: conservative pilot ratings, and lower values of pilot lead and gain giving rise to lower aileron activity and larger performance errors.

DESCRIPTION OF THE EXPERIMENT

This experiment was devised to take advantage of a simulation program already in progress on the Northrop Large Amplitude 3-Axis Flight Simulator. The program on the simulator was designed to compare results of aircraft evaluation performed in flight to those obtained with the simulator. The simulated aircraft was the Cornell Aeronautical Laboratory variable stability T-33.

Cornell Aero. Lab. had conducted an in-flight lateral-directional handling qualities evaluation of many aircraft configurations using their variable stability T-33. Although the evaluation of each aircraft consisted of many phases during which the pilot could decide upon the fitness of the configuration for air-to-air combat, the only part of interest here concerns a bank angle tracking task. The vertical needle of the attitude-direction indicator in the cockpit was used to provide a bank angle error signal. The pilot attempted to follow a program of step changes in bank angle. The commanded bank angles were either 30, 45 or 60 degrees and the train of steps ran for 100 seconds before repeating itself. An error in the commanded bank angle was presented to the pilot by a deflection of the ADI needle, (full deflection either right or left from zero representing 30°). When the needle was centered, the desired bank angle was being maintained.

It should be noted that the longitudinal dynamics of the aircraft were held constant for each configuration. Longitudinal characteristics which were suitable for a fighter were selected by Cornell. Then only parameters affecting lateral handling qualities were varied.

In order to perform the comparison of pilot transfer functions, the bank angle tracking task was modeled as in figure 1. A pilot parameter identification scheme

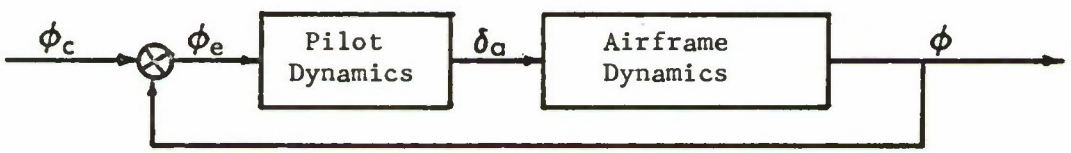


Figure 1. Pilot-Airframe Model

was used which required the assumption that the above pilot dynamics were of the form $K\frac{\tau s + 1}{s + b}$ where b is known.

The pilot parameter identification scheme was an analog method which used values of ϕ_e and δ_a to adjust the dynamics of two models to most closely match the speed and magnitude of the pilots response, (δ_a). Values of K and τ , the pilot's lead and gain, may be obtained with less than 30 seconds of a "good" signal. A "good" signal is one with both low (less than 1 cps) and high (more than 2 or 3 cps) frequency components. Since the frequencies required are normally presented in a pilot's output, system performance was adequate.

Accuracy of the identification circuit for purposes of determining the absolute value of parameters may be questionable. The circuit has properties which make the values of K linearly dependent upon the selected value of b . The value of τ is somewhat less influenced by the value of b , but the dependence is still heavy. Experiments with the circuit showed that a change in value of b from 4 to 6, (where the value of the transfer function being tracked was 4), resulted in a

+15% error in the value of τ . However, since we are here concerned with comparison of the same set of aircraft dynamics flown with and without motion cues, no attempt will be made to account for changes in the value of pilot lag which may have occurred. Instead, the data must be interpreted in light of these factors, as a comparison only.

Data runs were made by taking the configurations in random order and performing first a simulation with motion, and then one with no motion. In no case was the same configuration run consecutively using motion and then no motion.

The values of K, τ, $\int \phi_e dt$, $\int \delta_a dt$, and $\int \delta_e dt$ were recorded on a paper recorder for each run. Values of K and τ tended to rise to a relatively constant magnitude during each run. Indeed, in several cases the smooth shape of the τ curve suggested a learning curve as the pilot became more proficient at the task.

The parameters were read from the paper recorder at points satisfying the following criteria. A section of record was selected during which the pilot was very close to the correct bank angle, but where there was a "reasonable" amount of aileron activity. He was then controlling essentially the Dutch roll of the airplane. This provided a very stable signal from the parameter identification circuit.

Table 1, below, shows the configurations for which data was obtained. The aileron sensitivity, $L'_{\delta as}$, was adjustable by the pilot, but parameter identification used the actual aileron deflection. Hence, the gain values reflect the overall gain of degrees aileron to degrees error. The selected sensitivity is shown in the table below.

ω_d rad/sec	ζ_d	$\frac{N'_{\delta as}}{L'_{\delta as}}$	$L'_{\delta as}$ (deg/sec^2)/in	Cornell Rating
2.50	.090	.037	171	5.5
2.50	.090	-.034	154	4.0
2.50	.090	.059	140	7.0
2.44	.081	.037	54.1	5.5
2.44	.081	0.0	150	6.0

Table 1. Tested Configurations

RESULTS

Figures 2 through 7 on the following page indicates graphically a comparison of the results of the fixed and motion base simulations. Figure 2 shows that pilot ratings for fixed based simulation were more conservative than those from the simulation with motion cues. Figure 3 gives some indication that the pilot adapted a higher value of lead with the motion cues, although the data is quite scattered. Use of a higher gain in the aileron closures during moving base simulation is shown by Figure 4. Figure 5 indicates an increase in aileron activity when motion cues were present. Degraded performance without the motion cues is indicated by Figure 6. Figure 7 shows that elevator activity, a secondary control task, increased during the fixed base simulation. To summarize, we have associated with fixed base simulation: conservative pilot ratings, and lower values of pilot lead and gain giving use to lower aileron activity and degraded performance. The fixed base experiments also gave rise to larger elevator activity.

Several interesting possibilities suggest themselves. It seems evident that the pilot is able to use the motion to improve his performance. Part of the effect is in the ability of the vestibular apparatus to detect the acceleration present, and hence, provide a faster response. Part may be simply the inertial properties of the stick and the pilot's arm tending to resist the rolling of the aircraft, and hence, acting to reduce pilot time delay.

In interpreting Figure 3, the dependence of τ upon the selected value of b should be considered. The value for b was held constant during the experiment. However, one might expect the actual value of b in the pilot transfer function to be lower when motion cues were present. Hence, the value of b in the circuit may have been too large with motion, resulting in overestimation of τ for the moving base simulation. Therefore, one could interpret Figure 3 as an increase in lead or a decrease in lag. Further experimentation using multiloop models and parameter identification will be necessary in order to fully appreciate the significance of motion in pilot dynamics.

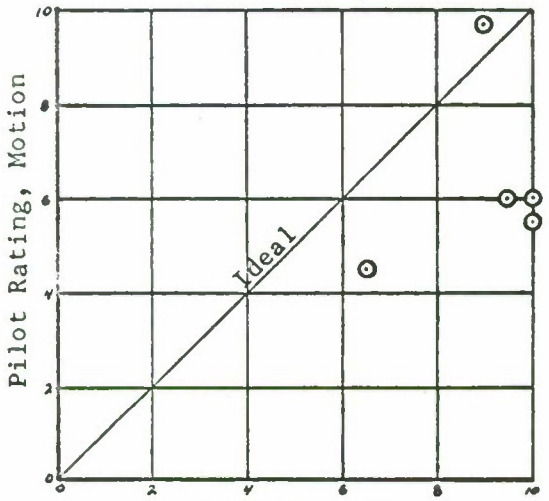


Figure 2, Pilot Ratings

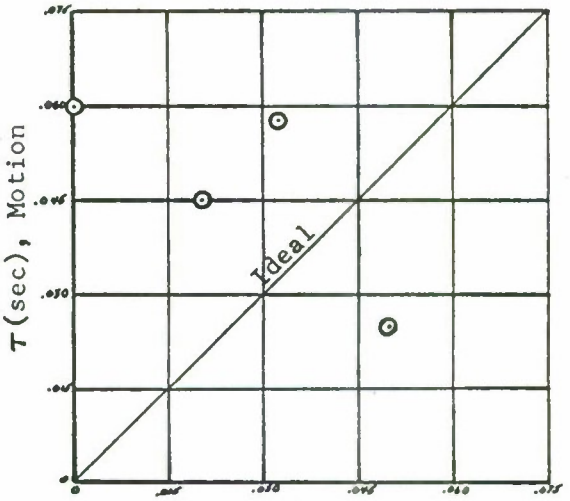


Figure 3, Pilot Lead

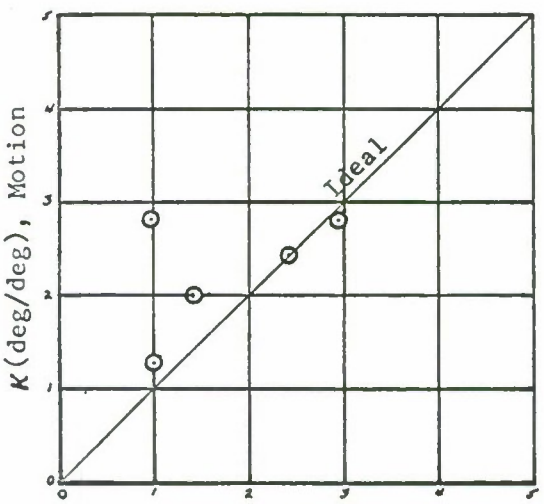


Figure 4, Pilot Gain

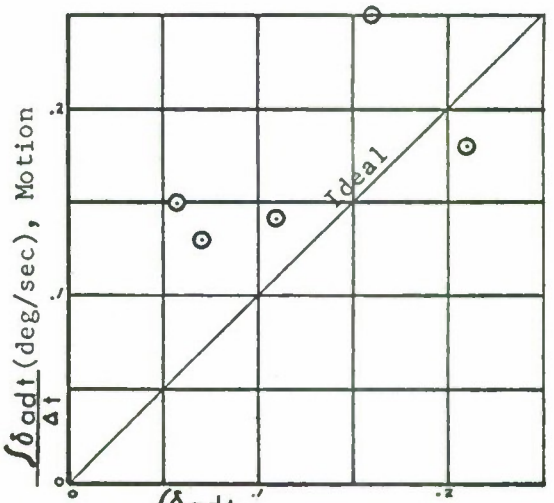


Figure 5, Aileron Activity

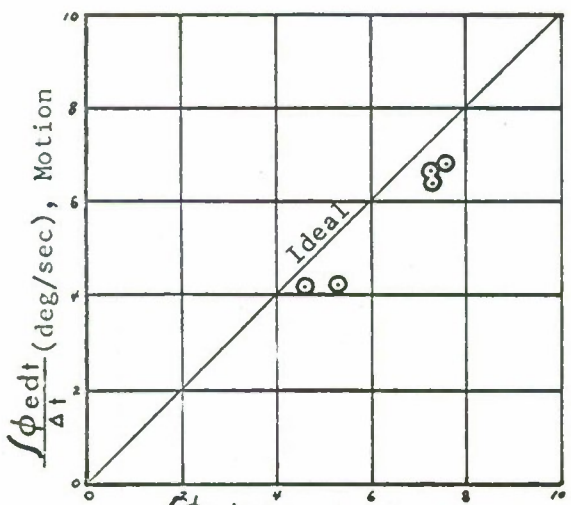


Figure 6, Error Rates

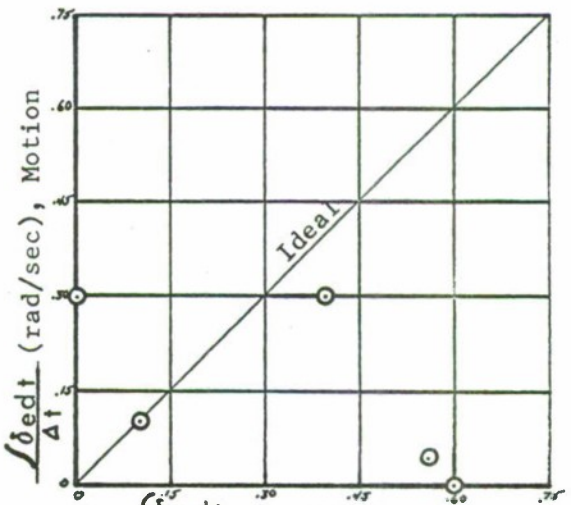


Figure 7, Elevator Activity

A SIMPLE FOURIER ANALYSIS TECHNIQUE FOR MEASURING THE DYNAMIC RESPONSE OF MANUAL CONTROL SYSTEMS*

R. Wade Allen and Henry R. Jex

Systems Technology, Inc.
Hawthorne, Calif.

Abstract

From a practical standpoint most dynamic response analysis techniques are complicated, expensive, and time consuming to mechanize and operate. These techniques usually require digital or special purpose analog equipment, and the computations are often performed sometime after data is collected. Adaptive parameter tracking techniques partially overcome these faults; but they either have prescribed forms which is an undesirable restriction for basic research, or in the case of free form methods the computational requirements become excessive.

To overcome these drawbacks an on-line Fourier Analysis technique has been developed which deals only with the input and error signals in a manual feedback control system. Several sine waves are used for the input, and the on-line data measurement includes the error variance and a simple means for obtaining the sine and cosine transforms of the error signal at each of the input frequencies. From this relatively small amount of data, we show simple ways to compute the open- and closed-loop dynamic response at input frequencies, and the relative amount of linearly correlated power in the error signal.

This paper describes the rationale and theory for the technique, and discusses two methods for mechanizing it; one involving standard analog computer components, and one based on conventional electromechanical components. Finally, a typical application of this method is presented.

INTRODUCTION

Fourier techniques for dynamic response measurements are not new. Mechanical engineers have used them for decades, and the earliest measurements of human operator dynamic response by Tustin (Ref. 1) and Russell (Ref. 2) used these techniques (see Ref. 3 for review of these studies). Some of the most comprehensive human dynamic response measurements were made by McRuer, et al., (Ref. 4) who used watthour meters to calculate Fourier coefficients.

---

*This work was performed as part of Contract NAS2-4405 from the NASA-Ames Research Center, Man-Machine Integration Branch.

The attractiveness of Fourier methods stems from the capability they allow for making measurements on-line, as tracking data is generated. Most general purpose time domain and spectral analysis methods involve digital computer processing, which requires data recording, digitizing, and careful bookkeeping of scale factors, run numbers, etc. Adaptive parameter matching techniques can be mechanized to work on-line (Ref. 5), while data is being generated. These techniques typically have prescribed forms, however, and often in basic tracking research this is an undesirable restriction.

FOURIER MEASUREMENTS

A general purpose Fourier measurement technique can be set up to operate on-line as shown in Fig. 1. A sum of sine waves is used for the input, and both the sine and cosine of each input frequency component are generated. These sine and cosine waves are then used to calculate the Fourier coefficients of any signal in the tracking loop simply by mechanizing the finite sine and

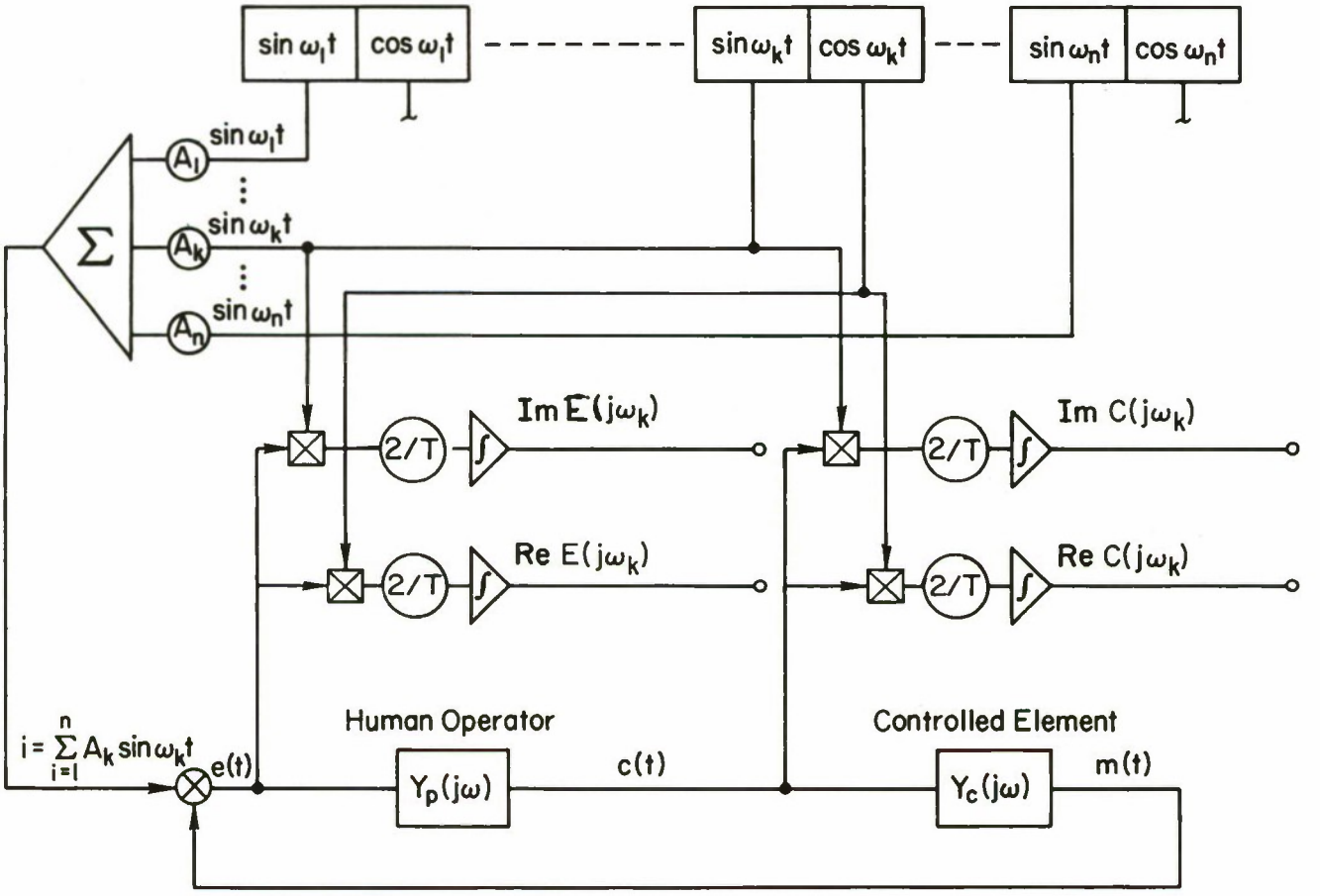


Figure 1. General Purpose Fourier Measurement Technique

cosine transforms* (Ref. 6),

$$F_c[x(t)]_k = \text{Re } X(j\omega_k) = \frac{2}{T}\int_{t_o}^{t_o+T} x(t)\cos(\omega_k t)dt ,$$

$$F_s[s(t)]_k = \text{Im } X(j\omega_k) = \frac{2}{T}\int_{t_o}^{t_o+T} x(t)\sin(\omega_k t)dt \qquad (1)$$

$$\text{where} \qquad \frac{\omega_k}{2\pi}\cdot T = n_k \text{ (an integer)}^{\dagger}$$

The finite Fourier transform of x(t) at the input frequencies is given by

$$F[x(t)]_k = \frac{2}{T}\int_{t_o}^{t_o+T} x(t)e^{-j\omega_k t}dt = \frac{2}{T}\int_{t_o}^{t_o+T} x(t)[\cos\omega_k t - j\sin\omega_k t]dt \qquad (2)$$

Thus given the Fourier coefficients of any two signals we can compute the Fourier transform ratio or describing function relating them at the various input frequencies. For instance, for the human operator we have

$$Y_p(j\omega_k) = \frac{\text{Re } C(j\omega_k) - j\text{Im } C(j\omega_k)}{\text{Re } E(j\omega_k) - j\text{Im } E(j\omega_k)} \qquad (3)$$

or in terms of the more familiar gain and phase relationships

---

*We have chosen here to normalize F_s and F_c with respect to time so that the resulting transform corresponds to the amplitudes of phase shifted sine waves, which when summed will give the original time signal x(t). Also the measurement interval T begins at an arbitrary time t_o which allows the input sine wave phasing to be arbitrary.

†This condition is theoretically required so that the input frequencies are orthogonal to each other, and the measurement process at one frequency is not influenced by effects due to the remaining frequencies (see Ref. 7, pp. 195-203 for a discussion of the orthogonality condition). As a practical matter, for high frequency inputs where the measurement period includes a large number of cycles (i.e., >100), the integer number condition is not critical.

$$\left|Y_{p_k}\right|_{dB} = 10 \log_{10} \frac{|C(j\omega_k)|^2}{|E(j\omega_k)|^2}$$

$$\measuredangle Y_{p_k} = \frac{180}{\pi}\left[\tan^{-1} \frac{\mathbf{Im}\, E(j\omega_k)}{\mathbf{Re}\, E(j\omega_k)} - \tan^{-1} \frac{\mathbf{Im}\, C(j\omega_k)}{\mathbf{Re}\, C(j\omega_k)}\right] (\text{degrees}) \qquad (4)$$

A SIMPLIFIED FOURIER TECHNIQUE

In order to obtain a describing function in the above manner, four multiplying and integrating elements are required for each frequency.* There is a simple way of limiting the required operations to just two multiplications and integrations per frequency, however, which essentially halves the amount of equipment and data required to obtain dynamic response measurements. This method involves computing the error-to-input describing function at each frequency, $Y_{ie}(j\omega_k)$, which is the ratio of the Fourier transforms $E(j\omega_k)$ to $I(j\omega_k)$. Writing this in polar form:

$$Y_{ie}(j\omega_k) = \frac{E(j\omega_k)}{I(j\omega_k)} = \frac{|E_k|e^{j\varphi_k}}{|I_k|e^{j\psi_k}} = \frac{|E_k|}{|I_k|} e^{j(\varphi_k - \psi_k)} \qquad (5)$$

It is helpful to visualize these quantities as "phasor" vectors, rotating at an angular frequency ω_k, as shown in Fig. 2. From this interpretation, it can be seen that the polar form of $Y_{ie}(j\omega_k)$ is given by the ratio of lengths of the i and e phasors and the net phase angle between them, $\measuredangle Y_{ie}$. By using the sine and cosine components of each generated input sinusoid as the (unity-amplitude) multipliers of e(t) in the error Fourier integrals, it can be shown that an "input-referenced Fourier transform," $\tilde{E}(j\omega_k)$ results, where $\tilde{E} = |E(j\omega_k)|\exp[j\measuredangle Y_{ie}]$. $\tilde{E}(j\omega_k)$ has the same Fourier <u>amplitude</u> components as $E(j\omega_k)$ but its <u>phases</u> (and hence real and imaginary components) are referenced to a unit vector lying on the i-phasor, as shown in Fig. 2, thereby giving the desired $\measuredangle Y_{ie}$ directly.†

*The watthour meter combines the multiplying the integration functions (see Ref. 12 for a discussion of this technique).

†For example consider the case where the input is a pure cosine wave, with $\psi_k = 0$. Then the Fourier transforms give ReE and ImE, and $\varphi_k \equiv \measuredangle Y_{ie}$, as noted. Provided the real (cosine) component of $I(j\omega_k)$ is used as the cosine multiplier for e(t), this same relationship will exist regardless of the arbitrary phase origin of $I(j\omega_k)$.

$\tilde{E}(j\omega_k)$ is measured by mechanizing the equations:

$$\text{Re}\ \tilde{E}(j\omega_k) \equiv \frac{2}{T}\int_{t_o}^{t_o+T} e(t)\cdot\cos(\omega_k t+\psi_k)dt$$

$$\text{Im}\ \tilde{E}(j\omega_k) \equiv \frac{2}{T}\int_{t_o}^{t_o+T} e(t)\cdot\sin(\omega_k t+\psi_k)dt \qquad (6a)$$

Noting that $i(t) = \sum_{k=1}^{n} A_k(\cos\omega_k t+\psi_k)$, the corresponding input components are:

$$\text{Re}\ \tilde{I}(j\omega_k) = \frac{2}{T}\int_{t_o}^{t_o+T} A_k\cos^2(\omega_k t+\psi_k)dt = A_k$$

$$\text{Im}\ \tilde{I}(j\omega_k) = \frac{2}{T}\int_{t_o}^{t_o+T} A_k\cos(\omega_k t+\psi_k)\cdot\sin(\omega_k+\psi_k)dt = 0 \qquad (6b)$$

That is, the input-referenced transform of the input consists of the real components A_k. Thus the magnitudes, $|I_k|$, required in Eq. 5 are merely the amplitudes of the input sinusoids, which are known a priori for on-line computations.

Having measured $Y_{ie}(j\omega_k)$ in the above manner we can now derive other describing functions of interest in analyzing manual control systems.

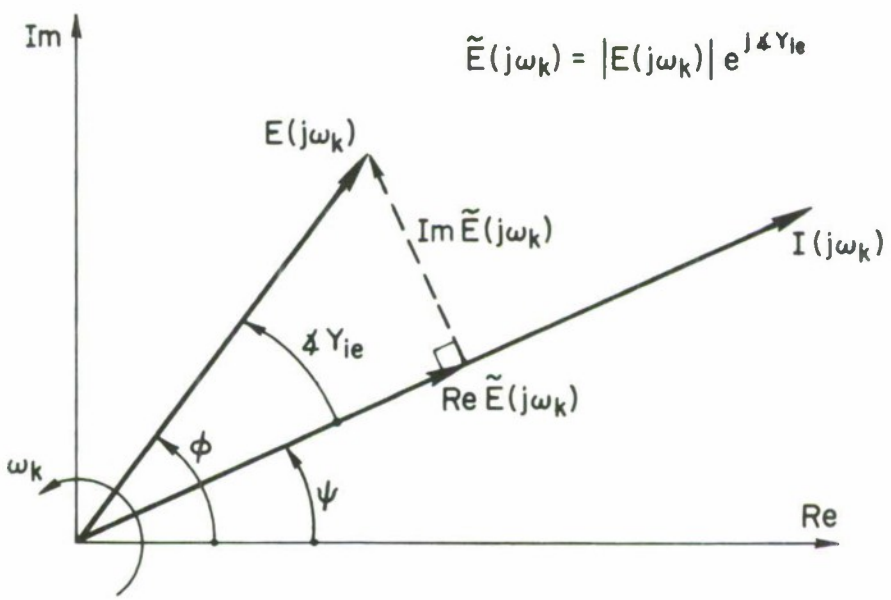


Figure 2. Error and Input Sinusoid "Phasor" Vectors

The open- and closed-loop Fourier transfer functions of a feedback system are related by simple vector equations. Thus once the error-to-input transform is obtained the open-loop transform $M(j\omega_k)/E(j\omega_k)$ is easily calculated with the relationship

$$Y_{OL}(j\omega_k) = \frac{M(j\omega_k)}{E(j\omega_k)} = Y_{ie}^{-1}(j\omega_k) - 1 \tag{7}$$

and the closed-loop transform is given by

$$Y_{CL}(j\omega_k) = \frac{M(j\omega_k)}{I(j\omega_k)} = 1 - Y_{ie}(j\omega_k) \tag{8}$$

If the controlled element transfer function $Y_c(j\omega)$ is known then the human operator's describing function can be calculated from the relationship

$$Y_p(j\omega_k) = \frac{Y_{OL}(j\omega_k)}{Y_c(j\omega_k)} \tag{9}$$

Given the transform of the error signal we can sum the squared magnitudes to obtain the input correlated error variance:

$$\overline{e_i^2} = \frac{1}{2}\sum_{k=1}^{n}\left[\mathrm{Re}^2\,\tilde{E}(j\omega_k) + \mathrm{Im}^2\,\tilde{E}(j\omega_k)\right] \tag{10}$$

Since the total error variance can be partitioned into input correlated and uncorrelated or remnant portions,

$$\sigma_e^2 = \overline{e_i^2} + \sigma_{e_n}^2 \tag{11}$$

merely by calculating σ_e^2 and the error transform we obtain a simple measure of remnant.

Care should be taken in measuring the error variance. Many times the error signal will have a significant offset or non-zero mean. This can result from human operator behavior, such as a pilot who consistently flies on the high side of a landing beam, or offsets may be due to simulation artifacts such as offset voltages in operational amplifiers. Thus in addition to measuring the mean square error power

$$\overbrace{e^2(t)} = \frac{1}{T}\int_o^T e^2(t)dt \qquad (12)$$

the mean value of the error signal

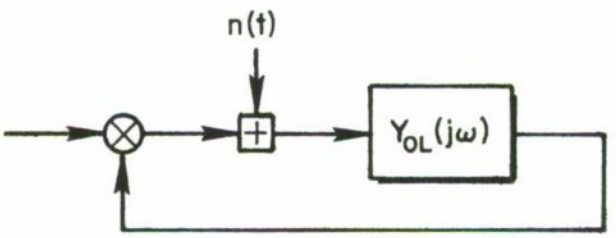
$$\overbrace{e(t)} = \frac{1}{T}\int_o^T e(t)dt \qquad (13)$$

should also be obtained. Then the error variance is given by the expression

$$\sigma_e^2 = \overbrace{e^2(t)} - \left[\overbrace{e(t)}\right]^2 \qquad (14)$$

MEASUREMENT ERRORS DUE TO REMNANT

The human operator injects remnant into the manual control loop. Recent research (Refs. 8 and 9) has shown that remnant measurements are best interpreted as a remnant source injected at the operator's perceptual input, in this case, the error signal as illustrated below. Because of this remnant



Manual Control Loop with Injected Remnant

the dynamic response measurements we make will be in error. In order to analyze this error let us consider the "black box" measurement shown in Fig. 3. The input $i(t)$ and its Fourier transform $I(j\omega)$ are perfectly known. Since we are attempting to identify the dynamic response $Y_{ie}(j\omega)$ at the input frequencies ω_k we require an accurate estimate of the error signal Fourier transform. The error signal consists of a portion linearly correlated with the input through the dynamic response function $Y_{ie}(j\omega)$ but it is also corrupted by the noise source $e_n(t)$ which corresponds to the closed-loop error remnant in the manual control loop.

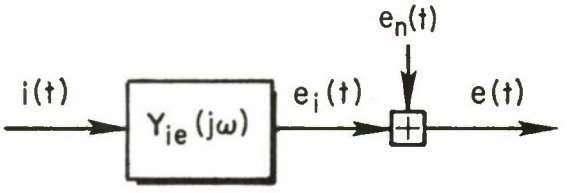


Figure 3. Equivalent "Black Box" Measurement Problem

The total error signal is given by

$$e(t) = e_i(t) + e_n(t) \tag{15}$$

with finite Fourier transform

$$\mathbf{F}[e(t)] = \frac{2}{T}\int_{t_o}^{t_o+T} [e_i(t) + e_n(t)]e^{-j\omega_m t}\, dt \tag{16}$$

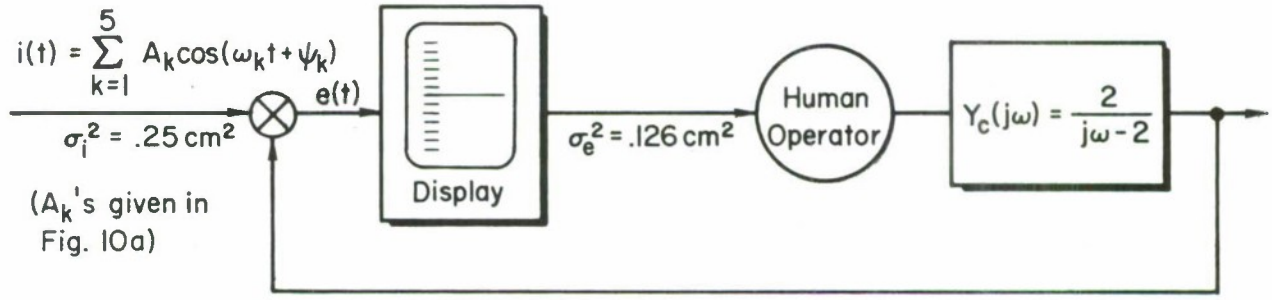
$$\text{where } \omega_m \doteq \frac{2\pi m}{T} \text{ and } m = 1,2,\cdots$$

The transform of $e_i(t)$ is desired but this signal is corrupted by the noise source $e_n(t)$ so that we can only obtain estimates of $F[e_i(t)]_k$,

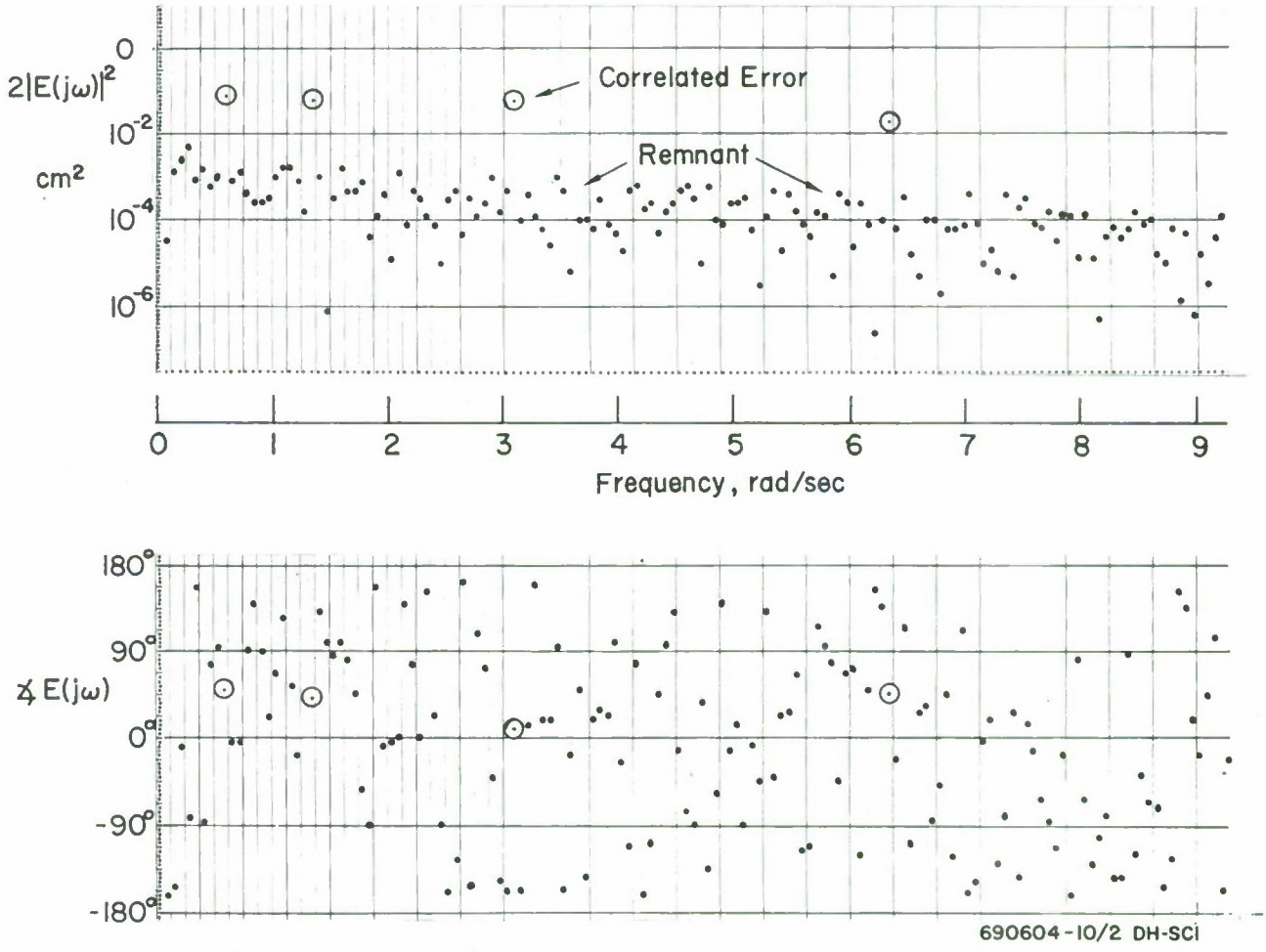
$$\hat{\mathbf{F}}[e_i(t)]_k = \frac{2}{T}\int_{t_o}^{t_o+T} [e_i(t) + e_n(t)]e^{-j\omega_m t}\, dt$$

$$= \mathbf{F}[e_i(t)] + \mathbf{F}[e_n(t)] \tag{17}$$

We see that this estimate is corrupted by the transform of the noise source, so the question arises as to the nature of $F[e_n(t)]$. A portion of the digitally computed finite Fourier transform of typical tracking data is shown in Fig. 4. The input for this experiment was a sum of cosine waves. The lowest four input sinusoid frequencies show up clearly as line spectra in the error power spectrum. In between the input correlated spikes in the error spectrum we see a relatively continuous remnant process, without any noticeable peaks. The error phase spectrum shows that the remnant spectrum has relatively random phasing.



a) Experiment



b) Error Signal Amplitude and Phase Spectra

Figure 4. Finite Fourier Transform of a Typical Tracking Error Signal for a Sum of Sinewaves Input

From the evidence given in Fig. 4 we conclude that $\mathbf{F}[e_n(t)]$ has a continuous amplitude spectrum with random phasing. The effect of this noise on the error transform at 5 input measurement frequencies is shown in Fig. 5 for some typical data. $\mathbf{F}[e_n(t)]$ causes a circular "area of uncertainty" in the measurement of the vector $\mathbf{F}[e_i(t)]$. The radii of the circular areas shown in Fig. 5 are the average magnitude of the remnant Fourier transform $|\mathbf{F}[e_n(t)]|$ at the measurement frequencies. The expected value of the estimate $\hat{\mathbf{F}}[e_i(t)]$ is the desired vector $\mathbf{F}[e_i(t)]$ so that the estimate is unbiased.

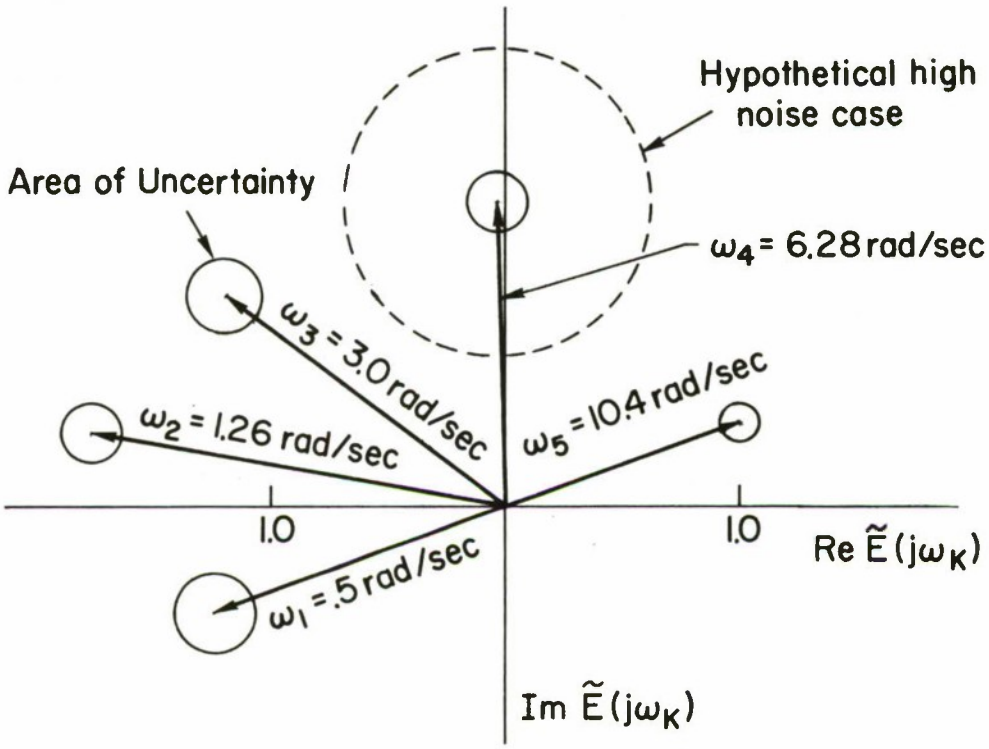


Figure 5. Error Fourier Transform Vectors and Noise Uncertainty Regions

If the estimates of the error-to-input describing function are noisy what are the resulting errors in the derived quantities, the open- and closed-loop dynamic response ratios, given by Eqs. 7 and 8? In the case of the closed-loop describing function Eq. 8 shows that $Y_{CL}(j\omega_k)$ has the same circular error region as the error to input dynamics $Y_{ie}(j\omega_k)$, since Y_{CL} is a linear transformation of Y_{ie}. This is not true for the open-loop dynamics, however, as shown by Eq. 7. In this case the derived quantity $Y_{OL}(j\omega_k)$ is a function of the inverse of the measured data $Y_{ie}(j\omega_k)$. An example of the error in Y_{OL} due to the circular error region of Y_{ie} is shown in Fig. 6 for data from Fig. 5. The actual data error region is reasonably circular. However if we hypothesize Y_{ie} data with a noise magnitude on the order of $1/2|Y_{ie}|$ we can see that the Y_{OL} error region is not symmetrically distributed. Thus low signal-to-noise conditions will cause biased estimates of Y_{OL}.

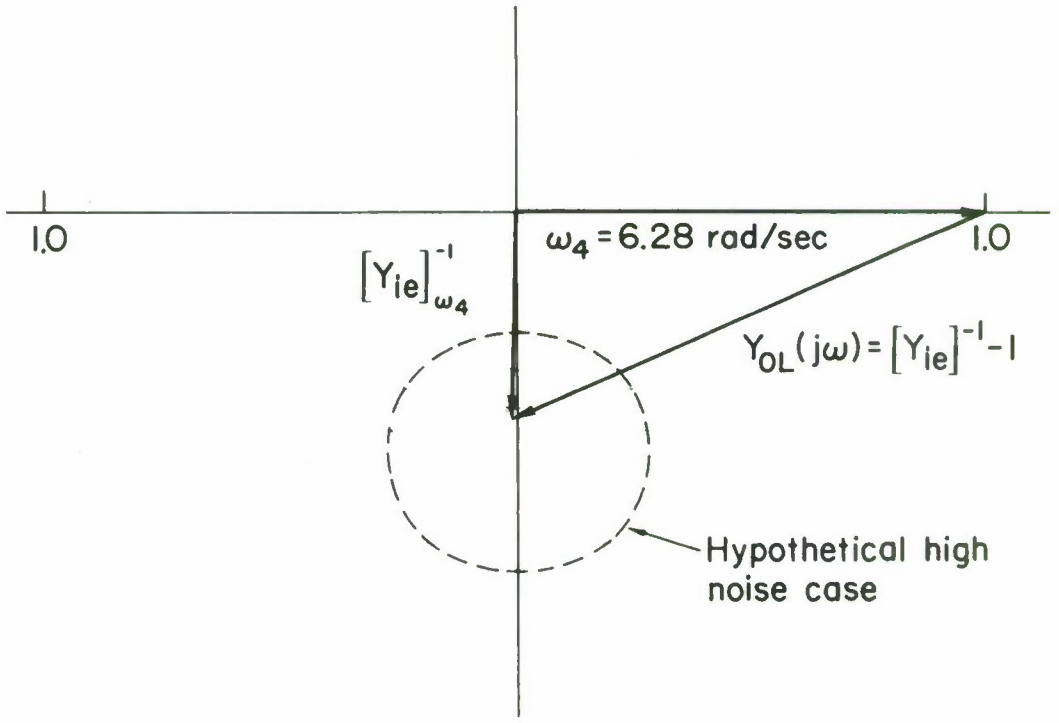


Figure 6. Open-Loop Response Calculation Near Crossover Frequency Showing Biased Noise Uncertainty Region

In order to minimize biased Y_{OL} measurements, and in general to achieve low variability data, we must achieve high signal to noise ratios in the error spectrum. One solution is to concentrate the correlated power into a relatively few number of sine waves as was the case for the data in Fig. 4. Another solution is to increase the measurement period. This reduces the remnant magnitude because the total remnant power will be partitioned over more Fourier transform harmonics. Finally we can merely make repeated measures of the error to input describing function and then calculate the mean of Y_{ie} before transforming to Y_{OL}.

RATIONALE FOR MEASURING THE ERROR-TO-INPUT DESCRIBING FUNCTION

The human operator's output signal or that of the controlled element (Fig. 1) might have been used for the present one-signal analysis technique. However, the following properties of the error signal, its spectrum and the error-to-input describing function make selection of e(t) advantageous from a practical standpoint for a wide variety of applications:

1. The correlated error spectrum [i.e., $\tilde{E}(j\omega_k)$] usually has the least dynamic range of any tracking signal in the loop. For the data given in Fig. 4 $\tilde{E}(j\omega_k)$ has nearly equal amplitude at each measurement frequency. Thus the dynamic range of the Fourier computations is minimized which eases the equipment accuracy requirements for mechanizing this technique.

2. The error-to-input describing function is the best overall dynamic response ratio to measure if it is also desired to derive the open- and closed-loop describing functions Y_{OL} and Y_{CL}. If the human operator output, c(t), to loop input, i(t), describing function is measured, then the controlled element dynamics $Y_c(j\omega_k)$ must be known accurately before Y_{OL} and Y_{CL} can be derived. In some applications an accurate description of Y_c is difficult to obtain. If the closed-loop describing function is measured, there may be inherent computational inaccuracies in deriving Y_{OL} and Y_{ie} since Y_{CL} approaches unity magnitude and zero phase shift at low frequencies for most systems.

3. With the present technique a measure of remnant is easily obtained for the error signal (Eqs. 10 and 11). This measure is relevant to recent remnant models for the human operator (Ref. 8), in which remnant is referred to the operator's input and depends on properties of the error signal in compensatory control situations.

4. The shape and properties of the error spectrum are often of direct interest for estimating man-machine performance (e.g., probability of exceeding tolerances, etc.).

MECHANIZATION

The Fourier transform technique described above has been mechanized in two different ways for use in measuring human operator dynamics.

One application which is easily mechanized on an analog computer is shown in Fig. 7. The sine and cosine waves are generated with neutrally stable feedback circuits. The error signal drives servo multipliers, and the sine and cosine multiplications are performed with the multiplying cups. This technique was mechanized on an EAI 1631R computer with type 16-7N servo multipliers. The phase shift in the measurement process was minimal even at 4π rad/sec. Amplitude instability in the sine wave generation circuits occurred at high frequencies ($\omega > 3$ rad/sec) but this was corrected with a nonlinear damping circuit.

The most recent mechanization of this describing function scheme is shown in Fig. 8.* Synchronous motors are used to drive sine and cosine potentiometers.

*This scheme is a variation of mechanical impedance techniques used by mechanical engineers. This particular mechanization was suggested by J. L. Durand of STI in September, 1965.

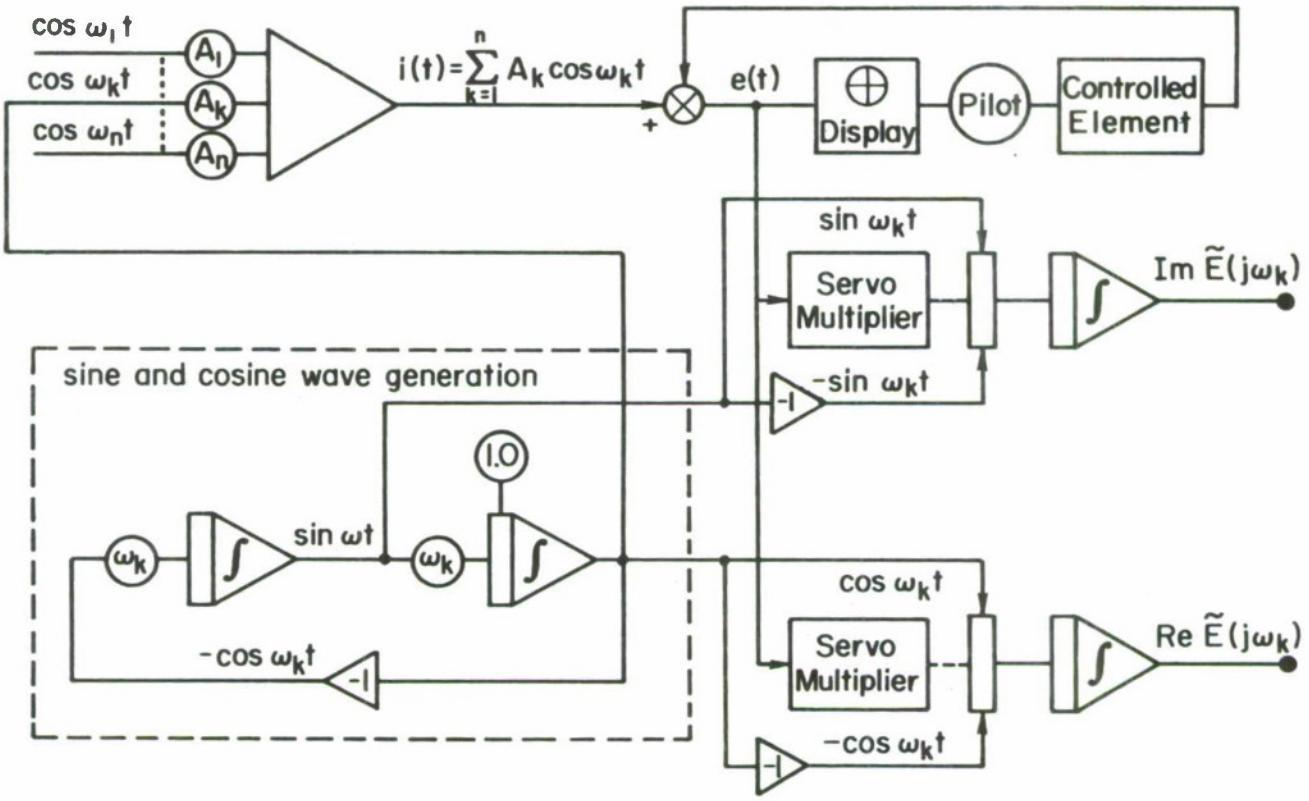


Figure 7. Analog Computer Mechanization of Simplified Fourier Analysis Technique

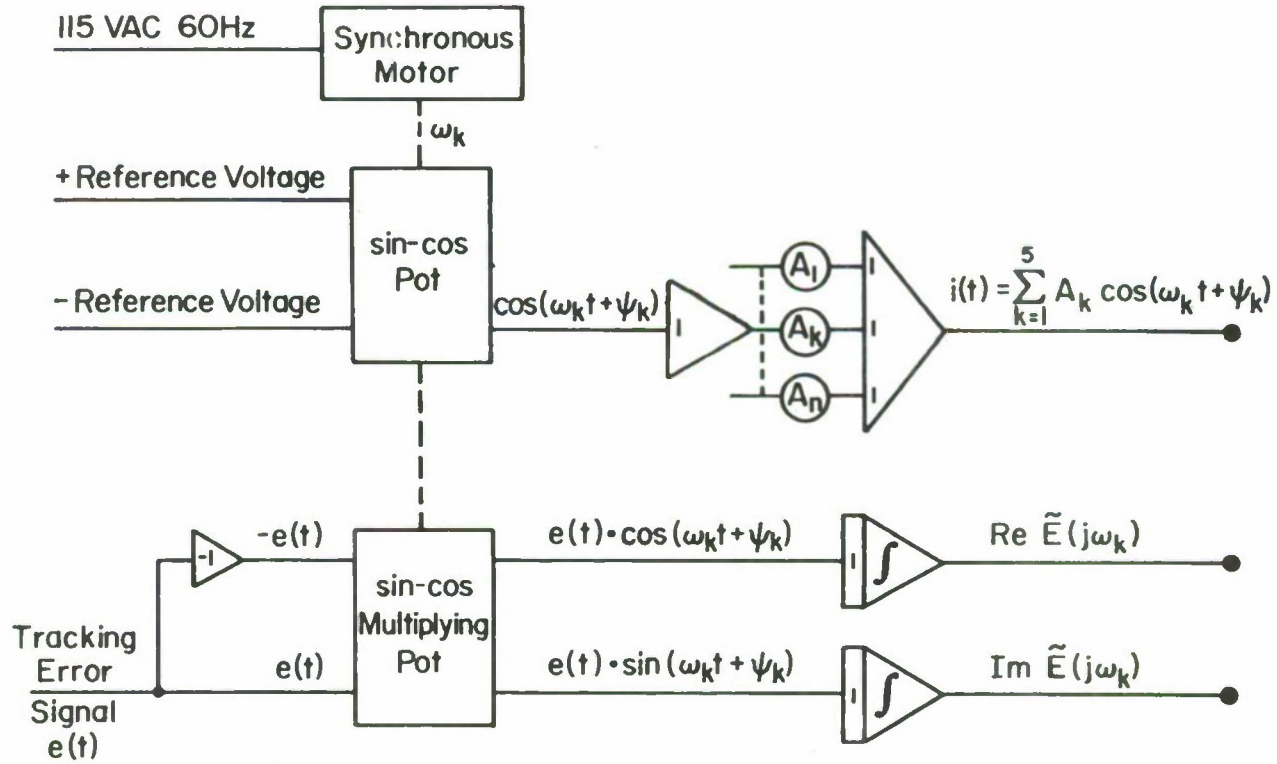


Figure 8. Single Channel of the Describing Function Analyzer

A reference voltage applied to one set of potentiometers yields cosine waves which are then summed to provide the input to the tracking task. Another set of sine-cosine multiplying potentiometers ganged to the input potentiometer shafts are used to perform the sine and cosine wave multiplication in the Fourier transform computation.

APPLICATION

The electromechanical technique was used very successfully in the experiment described in Ref. 10. Pilot describing functions for three repeated trials derived from the above technique are shown in Fig. 9. Contrary to some past experiments the describing function estimates are quite smooth across the spectrum and are stable from run to run. The performance measures show a similar stationarity.

Amplitude distributions and the error spectrum from the above experiment are shown in Fig. 10. While the input sum of 5 cosine waves has given a quite non-Gaussian amplitude distribution, the error signal which is composed of input correlated and remnant components is more normally distributed. The error power spectrum in Fig. 10c shows good measurement signal to noise levels at all five input frequency components.

In the practical application of both techniques described above some slight phase shift occurs in the measurement process at higher frequencies. This phase shift is computed by measuring the input Fourier transform $\tilde{I}(j\omega_k)$. $\tilde{I}(j\omega_k)$ should consist of entirely real components and any phase shifts are attributed to the measuring process and accounted for when computing the error-to-input describing function, $Y_{ie}(j\omega_k)$.

CONCLUDING REMARKS

The Fourier measurement technique described in this paper has proven simple to use in practice and has provided high production, low cost human operator describing functions on a routine basis. We have prepared a time-sharing computer program to compute $Y_{ie}(j\omega_k)$, $Y_{OL}(j\omega_k)$, $Y_p(j\omega_k)$, $Y_{CL}(j\omega_k)$, σ_e^2, $\overline{e_f^2}$, and $\sigma_{e_n}^2$ according to the formulas given previously, and reduced data can be produced within minutes after raw data is collected in the laboratory. Computation accuracies are comparable to those achieved in digital spectral analysis schemes, and are generally well within experimental research requirements.

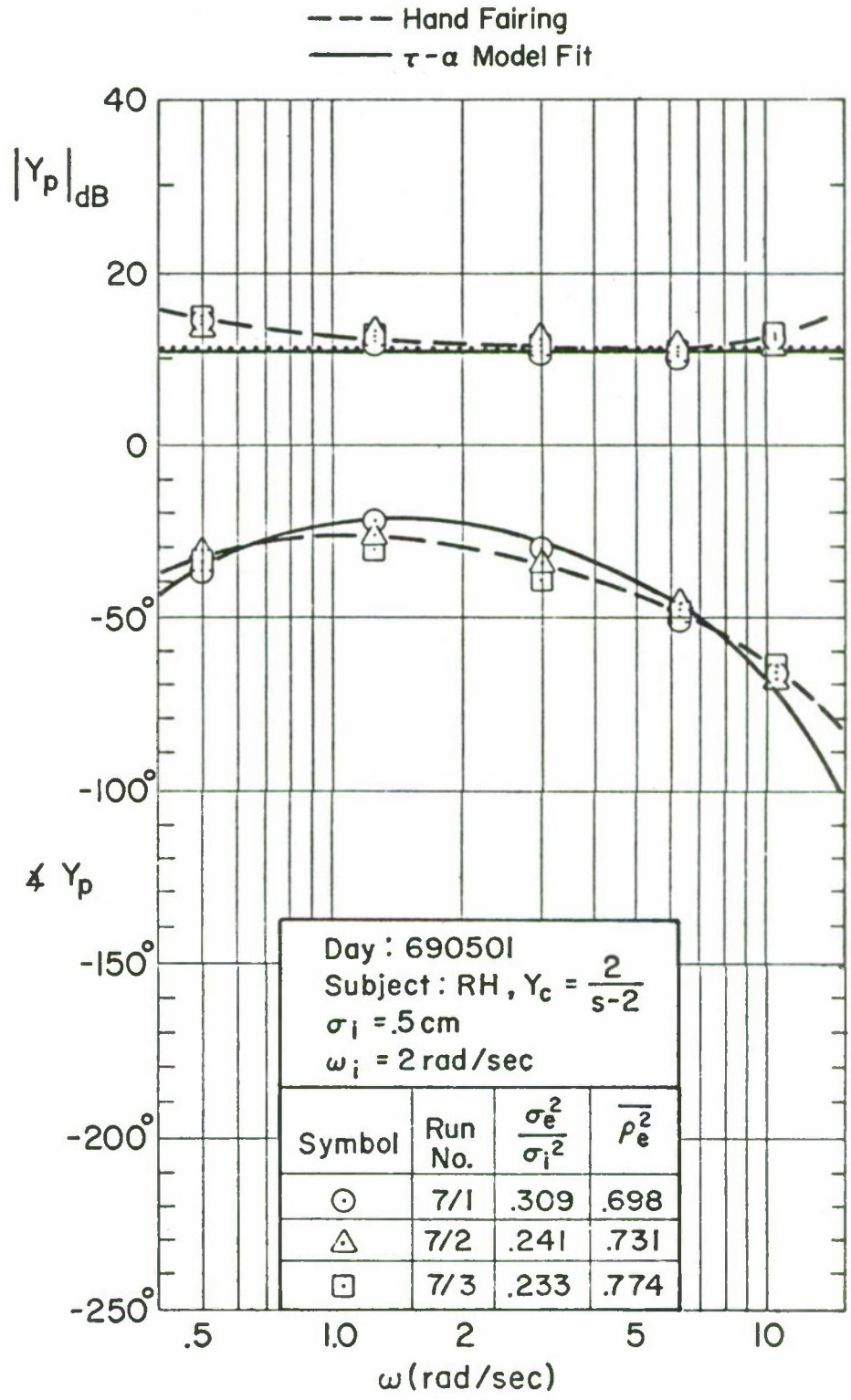
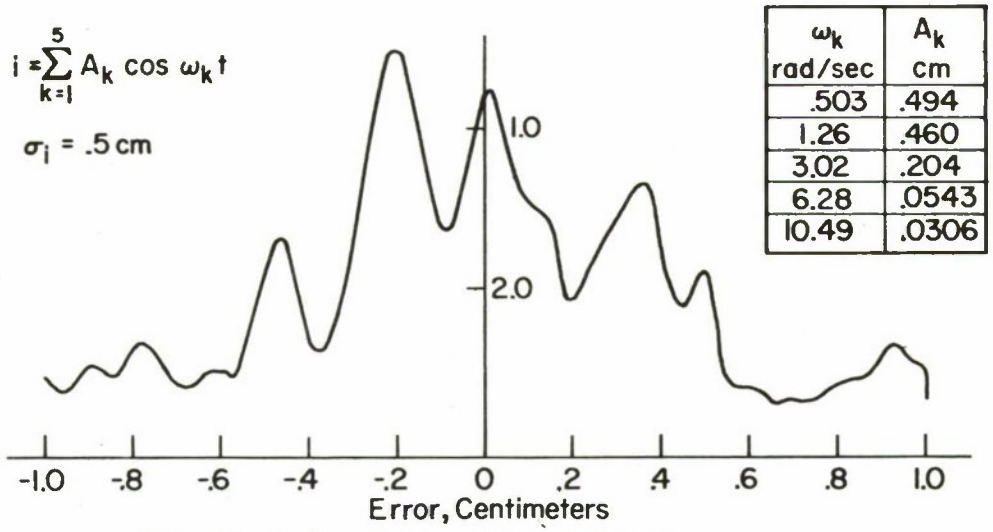
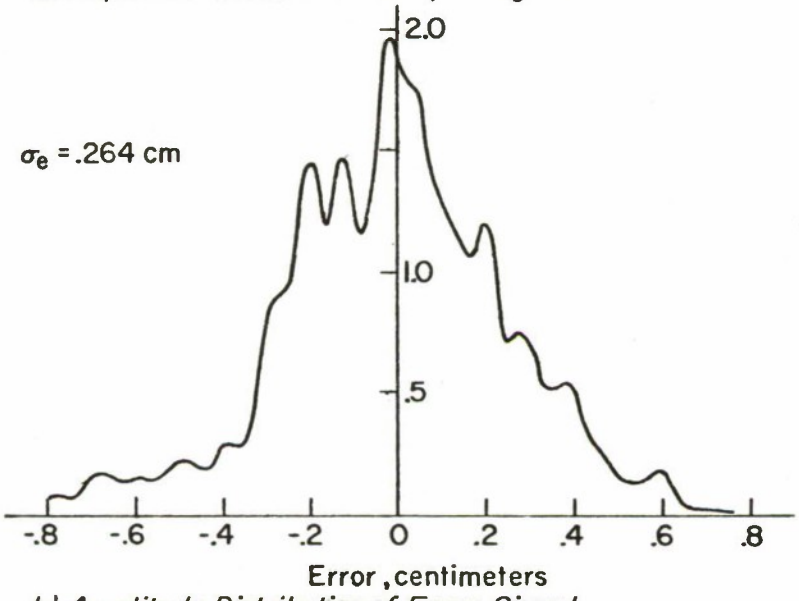


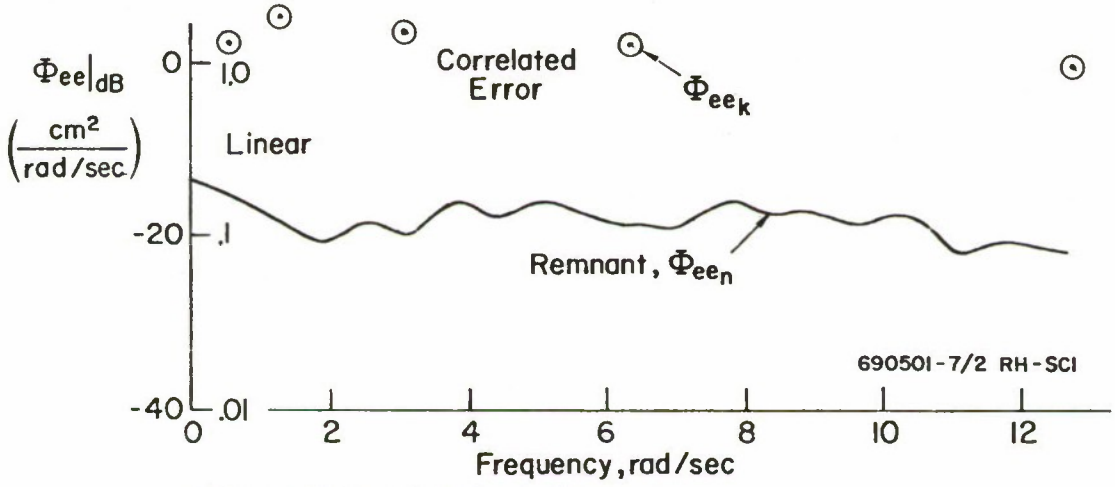
Figure 9. Three Repeated Describing Function Measurements Obtained with Simplified Fourier Analysis Technique



a) Amplitude Distribution of Input Signal



b) Amplitude Distribution of Error Signal



c) Error Signal Amplitude Spectrum

Figure 10. Input and Error Signal Characteristics for a Tracking Task with a Sum of Five Cosine Waves Input

REFERENCES

1. Tustin, A., "The Nature of the Operator's Response in Manual Control and Its Implications for Controller Design," J. IEE, Vol. 94, Part IIA, No. 2, 1947.

2. Russell, Lindsay, Characteristics of the Human as a Linear Servo-Element, M.S. Thesis, MIT, 18 May 1951.

3. McRuer, Duane T., and Ezra S. Krendel, Dynamic Response of Human Operators, WADC-TR-56-524, Oct. 1957.

4. 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.

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

6. Aseltine, John A., Transform Method in Linear System Analysis, McGraw-Hill, 1958.

7. Sokolnikoff, I. S., and R. M. Redheffer, Mathematics of Physics and Modern Engineering, McGraw-Hill, 1958.

8. Levison, W. H., S. Baron, and D. L. Kleinman, "A Model for Human Controller Remnant," Trans. IEEE, Vol. MMS-10, No. 4, Dec. 1969, pp. 101-107.

9. Jex, H. R., and R. E. Magdaleno, "Corroborative Data on Normalization of Human Operator Remnant," Trans. IEEE, Vol. MMS-10, No. 4, Dec. 1969, pp. 137-139.

10. Jex, H. R., and R. W. Allen, Validation of a Human Dynamic Response Test Battery, paper to be presented at the Sixth Annual Conference on Manual Control, Wright Patterson Air Force Base, Ohio, 7-9 Apr. 1970.

11. Seltzer, Lester J., and Duane T. McRuer, Survey of Analog Cross-Spectral Analyzers, WADC-TR-59-241, Dec. 1959.

A COMPARISON OF TECHNIQUES FOR MEASURING HUMAN OPERATOR FREQUENCY RESPONSE

Richard S. Shirley, NASA-Electronics Research Center

ABSTRACT

An experiment was performed to evaluate several techniques currently used for measuring the frequency response of pilots in compensatory tracking tasks. The same data were used to exercise each identification technique, permitting direct comparison of the results. Input-output data were obtained from human operators and known models. For the models noisy and noise-free data were obtained in open and closed loop situations. The techniques studied include frequency analysis, parameter tracking, filtering and cross correlation.

All methods provide good measurements in the region of crossover. Among those commonly used, the more computationally expensive techniques provide more accurate results away from crossover, despite showing greater variations in the presence of noise.

All methods deteriorate when signal levels are low. Closing the loop reduces error power, and the controlled dynamics $1/s$ or $1/s^2$ reduce low frequency error power, thus leading to poorer identification of known models than in the open loop case.

TABLE OF SYMBOLS

A_k	coefficient multiplying sinusoid of frequency ω_k to form the input
A_{ck}	Fourier sine coefficient at frequency ω_k of the control stick output, see equation 6
A_{cj}	same as A_{ck}, only at frequency ω_j
A_{ek}	same as A_{ck}, only of the error signal, e
A_{mk}	same as A_{ck}, only of the system output, m
ARC	NASA-Ames Research Center

B_{ck}	Fourier cosine coefficient at frequency ω_k of the control stick output, see equation 7
B_{cj}	same as B_{ck}, only at frequency ω_j
B_{ek}	same as B_{ck}, only of the error, e
B_{mk}	same as B_{ck}, only of the system output, m
BBN	Bolt Beranek and Newman, Inc., Cambridge, Mass.
c	control stick output, see figure 1
$\underline{C}$	vector $[c(n\Delta t)]$
CL	closed loop
CN	colored noise
e	error signal, see figure 1
$\underline{E}$	see equation 14
ERC	NASA-Electronics Research Center
$\underline{f_i}$	vector $[f_i(n\Delta t)]$
$\underline{F}$	matrix $[\underline{f_i}]$
$(Filt)_i$	i^{th} filter
FFT	fast Fourier transform
$h_p(t)$	the pilot's impulse response
i	input, see figure 1
i	integer subscript
j	integer subscript
k	integer subscript
K	gain
K_1	parameter of pilot model, see equation 11

K_2	parameter of pilot model, see equation 11
LRC	NASA-Langley Research Center
m	system output, see figure 1
m'	integer, number of filters
M	number of lags used in cross correlation
n	integer subscript, integer
N	integer number of data samples
NN	no noise
OL	open loop
p_i	i^{th} pole of a filter, see equation 10
PT	parameter tracker
s	the Laplace variable
t	time, seconds
T	warm-up period, 24 seconds
T_1	data-taking period, 216 seconds
WN	white noise
$x_{nn}(t)$	the pilot's remnant
XCOR	cross correlation
Y_c	vehicle dynamics, see figure 1
Y_m	measured dynamics
Y_p	pilot dynamics, see figure 1
$\underline{\beta}$	vector $\left[\beta_i\right]$
β_i	coefficient of the i^{th} filter output
Δt	time increment, .02 seconds

θ_k	phase shift of sinusoid of frequency ω_k
λ	time shift, seconds
σ_i	standard deviation of the input disturbance
τ	time shift, seconds
τ	parameter of pilot model, see equation 11
Φ_{ec}	cross power spectrum between e and c
Φ_{ee}	auto power spectrum of e
Φ_{nn}	auto power spectrum of the remnant
ω	frequency, rad/sec
ω_{co}	crossover frequency, rad/sec
ω_k	k^{th} input frequency, rad/sec
ω_j	j^{th} frequency between the input frequencies, rad/sec

BACKGROUND

Because today's vehicles are so complex, for both financial and human reasons it is necessary to analyze and predict man-machine system performance before building the machine for the man to test. Such analyses make use of prior experience, psychology, materials sciences, chemistry and other disciplines. In the last twenty years man-machine system analysis has also made use of control theory with rewarding results: one has only to look at the references of this paper to see many examples. In order to apply control techniques however, one must be able to model both the man and the machine. Modelling the machine has proven difficult, while modeling the man has proven nearly impossible. Men as servo elements are nonlinear, time-varying, adaptable and noisy. To make matters worse, men think. One may be glad his pilot thinks, but the engineer who tries to model him isn't so sure. Some progress has been made to simulate a pilot's mental processes by the introduction of the "Zeibolz controller" (references 26,27,31,48), optimal control theory (references 5,6,28) and other techniques, but even these

models are not yet sufficiently developed to be easily used by system designers. Despite a lack of a comprehensive model of the human operator, and despite the difficulties mentioned above, many models of the human operator have been developed which are valid in certain restricted but very useful cases. One of the more important cases is in either a pursuit or compensatory tracking task as shown in figure 1.

In figure 1 it is assumed that the man is acting primarily as a servo element, and that for the most part he is not "thinking", for as soon as he acts on information from outside the control loop the model of him is no longer valid. While keeping a plane on a glidepath, hovering a helicopter, driving a car or steering a ship, men often act in the capacity of servomechanisms in a compensatory system. Using a human being in this manner is reasonable, for in times of emergency his additional capabilities of pattern recognition, adaptability and decision-making (or "thinking") are available.

I. INTRODUCTION

When a dynamic control model of a pilot is developed, it must somehow undergo evaluation. In the case of a servomechanism one would measure the frequency response of the servo. In the case of the pilot this has not proven so simple. First Tustin (reference 44) and since then many others have considered the problem of measuring the frequency response of the human operator. The techniques used include various forms of cross correlation, regression analysis, frequency analysis and differential equation coefficient methods. The present paper describes an experiment performed to evaluate some of these techniques. Such an experiment is needed, for it is often difficult to compare the findings of different experimenters when they use different identification techniques.

II. THE EXPERIMENT

The following experiment was designed to evaluate several techniques for measuring the frequency response of the pilot in a compensatory tracking task. Of greatest importance was the requirement to use the same data in exercising each identification technique in order to permit a direct comparison of the results. Input-output data were obtained from human operators and known models.

For the known models noise-free and noisy data were obtained in both open loop and closed loop situations. The control configurations used are shown in figures 1 and 2.

It should be kept in mind that this experiment compared techniques on single runs, while normal practice is to present data which are an average of a number of runs.

A. The Identification Techniques Studied

The techniques for measuring pilot frequency response which were studied were of four basic types: frequency analysis, filtering, parameter trackers and cross correlation. The various implementations used for the experiment are briefly described and references below, while Table I gives a summary.

1) The method of Fourier coefficients (frequency analysis), NASA-Electronics Research Center (references 32,36).

The basic development of the technique was done by McRuer, et al, and is described in reference 32. The version used will be described in some detail because the experiment is built around it.

The method of Fourier coefficients utilizes a special form of input disturbance, or forcing function, to permit complete and accurate frequency response measurements to be obtained from a human being in a compensatory control loop as shown in figure 1. This is accomplished by concentrating the input power at a small number of discrete frequencies in order to maximize the signal to noise ratio at these frequencies, and to permit the direct measurement of noise levels between these frequencies. Specifically, the input disturbance is given by

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

where $\Delta t = .02$ seconds and n ranges from 1 to 10,800. The ω_k are spaced roughly evenly on a logarithmic scale between .1 and 30 rad/sec, and the A_k are alternated in sign to avoid high initial transients, and reduced in magnitude at the higher frequencies, as shown in Table II.

During the run data are taken every Δt seconds at the system error, $e(t)$, the control stick output, $c(t)$, and the system output, $m(t)$, to give the series $e(n\Delta t)$, $c(n\Delta t)$, and $m(n\Delta t)$. The linear describing functions for the human operator and for the open loop can then be calculated using

$$\angle Y_p(\omega_k) = \tan^{-1}\left|\frac{B_{ck}}{A_{ck}}\right| - \tan^{-1}\left|\frac{B_{ek}}{A_{ek}}\right| \tag{2}$$

$$\left|Y_p(\omega_k)\right| = \left|\frac{A_{ck}^2 + B_{ck}^2}{A_{ek}^2 + B_{ek}^2}\right|^{1/2} \tag{3}$$

$$\angle Y_p Y_c(\omega_k) = \tan^{-1}\left|\frac{B_{mk}}{A_{mk}}\right| - \tan^{-1}\left|\frac{B_{ek}}{A_{ek}}\right| \tag{4}$$

$$\left|Y_p Y_c(\omega_k)\right| = \left|\frac{A_{mk}^2 + B_{mk}^2}{A_{ek}^2 + B_{ek}^2}\right| \tag{5}$$

where

$$A_{ek} = \sum_{k=1}^{N} e(n\Delta t)\, \sin(\omega_k n\Delta t) \tag{6}$$

$$B_{ek} = \sum_{k=1}^{N} e(n\Delta t)\, \cos(\omega_k n\Delta t) \tag{7}$$

and A_{ck}, B_{ck}, A_{mk}, and B_{mk} are similarly defined. In addition the human operator's remnant (noise injected into the loop by the pilot) can be calculated at frequencies ω_j, between the input frequencies, ω_k, using

$$\Phi_{nn}(\omega_j) = \frac{K\left|Y_c(\omega_j)\right|^2}{\left|1 + Y_p Y_c(\omega_j)\right|^2} \left(A_{cj}^2 + B_{cj}^2\right) \qquad (8$$

where A_{cj} and B_{cj} are also defined similarly to A_{ek} and B_{ek} in equations 6 and 7.

The results of applying this identification technique to the input-output data $i(n\Delta t)$, $e(n\Delta t)$, $c(n\Delta t)$ and $m(n\Delta t)$, are the pilot and open loop frequency response curves (actually a series of points at frequencies ω_k), as well as the pilot's remnant. The corresponding equations are 2, 3, 4, 5, and 8.

The unique feature of the Fourier analysis program used at NASA-Electronics Research Center is that the results are available five seconds after the period of data-taking. The rapid processing of the data is obtained by using a modified fast Fourier transform (reference 36). The experiment is built around the technique for two reasons. First, the rapid processing of the data allowed each run to be immediately checked for obvious errors. Secondly, the hybrid computer used permitted the data to be readily processed by other techniques.

2) The method of Fourier coefficients (frequency analysis), Bolt Beranek and Newman, Inc. (reference 32).

The approach used by BBN is basically the same as that described in the previous section, except that a standard form of the FFT is used. The number of data points must be a power of two (BBN used 4096 points per run) and all the data must be available in core before processing can begin. In return for these restrictions all the Fourier coefficients (2048 of them) are obtained. The additional coefficients between the input sinusoid frequencies permit a more accurate determination of the remnant to be made, for a weighted sum of the noise measurements can be used instead of relying on a single measurement. In addition the remnant measurements used are closer to the input frequencies, and thus give a better indication of the signal to noise ratio at the input frequencies. This additional knowledge allows one to have a better feeling for the reliability of the describing function measurements.

Direct comparison of this technique with the others was not possible because of the restriction to 4096 data points. Consequently three special runs were made in which technique 1 was also restricted to 4096 data points, and Δt was .05 seconds. As expected, both

techniques gave identical results for the amplitude ratio and phase, but the remnant measurements differed as shown in figures 31 to 36.

3) Orthogonal filters, NASA-Ames Research Center (references 18,46,47).

Filtering is a second basic technique for identifying a system from its input-output relationship. The input, $e(t)$, is fed into a series of filters, $(Filt)_i$, yielding a set of filter outputs, $f_i(t)$. By weighting these filter outputs with coefficients β_i and adding the weighted outputs together, one can attempt to match the system output, $c(t)$. In the discrete case this can be accomplished in a least squares sense as follows: let

$e(n\Delta t)$ = the system input samples

$c(n\Delta t)$ = the system output samples

$f(n\Delta t)$ = the output samples of the i^{th} filter

β_i = the desired coefficients of the filter outputs

m' = the number of filters used

N = the number of data points (input-output pairs)

$\underline{C} = \{c(n\Delta t)\}$ (N x 1 vector)

$\underline{f_i} = \{f_i(n\Delta t)\}$ (N x 1 vector)

$\beta = \{\beta_i\}$ (m' x 1 vector)

$\underline{F} = \{\underline{f_i}\}$ (N x m' matrix)

If we set

$$\underline{F}\ \beta = \underline{C}$$

premultiply by $\underline{F}^T$

$$\underline{F}^T\ \underline{F}\ \beta = \underline{F}^T\ \underline{C}$$

and solve for $\underline{\beta}$

$$\underline{\beta} = (\underline{F}^T \underline{F})^{-1} \underline{F}^T \underline{C} \tag{9}$$

it can be shown that the $\underline{\beta}$ obtained minimizes the function

$$\sum_{n=1}^{N} \left[c(n\Delta t) - \beta_i f_i(n\Delta t)\right]^2$$

Once the filter coefficients are determined, calculation of the pilot's frequency response is an easy matter (reference 18).

It still remains to choose the filters. One can obtain orthogonal filters in the s-plane as follows:

$$(\text{Filt})_1 = \frac{\sqrt{2p_1}}{s + p_1}$$

$$(\text{Filt})_2 = \frac{\sqrt{2p_2}\ (s - p_1)}{(s + p_1)(s + p_2)}$$

$$(\text{Filt})_i = \frac{\sqrt{2p_i}\ \ (s - p_1)(s - p_2)\ldots(s - p_{i-1})}{(s + p_1)(s + p_2)\ldots(s + p_i)} \tag{10}$$

where s is the Laplace operator, and the p_i give the locations of the poles in the s-plane. The version used by Wingrove at ARC for this experiment has five poles located at s equal to -1, -2, -4, -8, and -16. Of the 10,800 data points taken, every other point was used for a total of N = 5,400.

A unique aspect of this technique is that the input-output

samples are shifted relative to each other by an amount roughly equal to the expected time delay of the pilot's transfer function. The shift reduces the effects of noise on the calculated values of the pilot's frequency response, thus leading to more accurate results (references 46,47).

4) Filtering and system realization, NASA-Electronics Research Center (references 20, 21, 37).

The basic approach of this technique is the same as the one described in 3) above, the main differences being the filters used and the fact that a system realization (and transfer function) is obtained in addition to the frequency response of the pilot. The filters are non-orthogonal poles in the s-plane as shown in Table III. As described in references 20, 21 and 37 a modified B.L. Ho procedure is used to obtain a system realization from the filter coefficients $\underline{\beta}$. In the case of noise-free data the system realization obtained is least-order, and the order need not be specified by the programmer. In the case of noisy data the desired system order must be specified, and the system obtained will be an approximation to the actual system.

5) Parameter tracker, NASA-Langley Research Center (references 1,2,3).

If one can assume a model for an unknown or time-varying system, the model can be programmed with variable parameters on an analog computer. The input to the system is fed into the model, and the difference between the system output and the model output is used in conjunction with a gradient technique to vary the adjustable parameters. With proper programming the parameters can be made to converge to values which minimize the difference between system and model outputs. Although the stability and the speed of convergence of such parameter trackers is of concern, the technique has the advantages of being physically easy and inexpensive to implement (requiring only an analog computer), as well as providing a time history of the parameters when the system being identified is time-varying.

The parameter tracker used by Adams at Langley assumes a three-parameter model of the pilot given by

$$Y_p(s) = \frac{\frac{K_1}{\tau}\left(1 + \frac{K_2}{\tau}s\right)}{\left(1 + \frac{1}{\tau}s\right)^2} \qquad (11$$

6) Parameter-tracking version of the crossover model, Oakland University (references 25,32).

The crossover frequency, ω_{co}, is that frequency at which the open loop amplitude ratio, which is the product of the absolute values of the pilot dynamics and of the vehicle dynamics, passes from greater than unity to less than unity. The crossover model states that in the region of crossover the open loop dynamics are closely approximated by

$$Y_p Y_c(s) = \frac{K e^{-\tau s}}{s} \qquad (12$$

This parameter tracker is thus a two-parameter model as developed by Jackson.

7) Cross correlation, NASA-Ames Research Center (references 46,47).

The linear input-output relationship for a pilot can be expressed by the concolution integral

$$c(t) = \int_0^T h_p(\tau)\, e(t - \tau)\, d\tau + x_{nn}(t) \qquad (13$$

where

c = the pilot's output

h = the pilot's linear impulse response

e = the pilot's input

x_{nn} = the pilot's remnant

The digital approximation is made as

$$c(n) = t \sum_{k=0}^{M} h_p(k)\, e(n - k) + x_{nn}(n) \qquad (14$$

or

$$\underline{c} = \underline{E}\,\underline{h_p} + \underline{x_{nn}}$$

where

$$\underline{E} = \Delta t \begin{bmatrix} e(n_o) & e(n_o-1) & \cdots & e(n_o-M) \\ e(n_o+1) & & & \vdots \\ \vdots & & & \vdots \\ e(N) & \cdots & \cdots & e(N-M) \end{bmatrix} \qquad (15$$

$$\underline{h_p} = \underline{h_p}(n)$$

$$\underline{c} = \left[c(n_o), c(n_o+1), \ldots, c(M)\right]^T$$

$$\underline{n} = \left[x_{nn}(n_o), x_{nn}(n_o+1), \ldots, x_{nn}(M)\right]^T$$

An estimate of $\underline{h_p}$ can be obtained (references 46,47) using

$$\hat{\underline{h}}_p = (\underline{E}^T \underline{E})^{-1} \underline{E}^T \underline{c} \qquad (16$$

Determination of the pilot's frequency response from $\hat{\underline{h}}_p$ is an easy step involving the use of the Fourier transform.

For the version used the input and output signals are shifted relative to one another by a time of λ seconds when there is a time delay between the input and output, leading to a better estimate of the system being identified. The delay used was .12 seconds when Y_c was 1/s, and .20 seconds when Y_c was $1/s^2$.

It is possible to calculate the pilot's frequency response at any frequency, but for clarity the results of the technique are shown in the figures as points. The same is true for technique 3.

B. The Computers Used

Technique 1 was implemented on a Beckman-2200 (analog) SDS-9300 (digital) hybrid computer. The analog portion has over 150 amplifiers, while the digital portion has a 32K core with a 1-1/2 micro-second cycle time. The computer and peripheral equipment was versatile

enough to have permitted the following:

1) For each run the dynamics to be controlled were simulated, data taken, and technique 1 was implemented. For each run this gave the amplitude ratio and phase for both Y_p and Y_pY_c, as shown in figure 1.

2) Simultaneously both parameter trackers (techniques 5 and 6) were implemented on the analog computer, providing on-line records of the time histories of Y_p and Y_pY_c.

3) At the end of each run the sampled data, $i(n\Delta t)$, $e(n\Delta t)$, $c(n\Delta t)$, and $m(n\Delta t)$ were sequenced and saved on digital magnetic tape. These tapes provided the data for techniques 2, 3, 4 and 7.

Technique 2 was implemented on an SDS-940 digital computer, while techniques 3, 4 and 7 were run on an IBM 360 digital computer.

C. The Dynamics Identified

The dynamics used to test the various identification techniques are summarized in Table IV, and include:

1) $Y_p(s) = 1/(s + 1)$; $\quad Y_p(s) = (1/(s + 1))^2$

These simple dynamics were used as shown in figure 2/a to calibrate technique 1.

2) $$Y_p(s) = 5\left[\frac{8.7 - s}{8.7 + s}\right]; \quad Y_c(s) = 1/s$$

As shown in figure 3, the dynamics given by Y_p above are an approximation of the frequency response of the human operator controlling the vehicle dynamics 1/s, or pure velocity control. Measurements were taken as shown in figure 2/b.

3) $$Y_p(s) = .75\left[\frac{8 - s}{8 + s}\right]\left[\frac{4.4s + 1}{.04s + 1}\right]\left[\frac{1}{.12s + 1}\right]; \quad Y_c(s) = 1/s^2$$

As shown in figure 4, the dynamics given by Y_p above approximate the frequency response of the human operator controlling the vehicle dynamics $1/s^2$, or pure acceleration control. Measurements were taken as shown in figure 2/b.

4) $Y_p(s)$ = a human pilot; $Y_c(s) = 1/s$ or $1/s^2$

A technique might perfectly identify known, fixed systems, but poorly identify the human operator. Consequently all the techniques studied were used to process data taken from a human operator in a compensatory tracking task controlling the dynamics $1/s$ and $1/s^2$.

D. The Disturbance Input

The input, $i(n\Delta t)$, in figures 1 and 2/b, was in every case the sum of sinusoids shown in Table II.

E. The Noise

It is generally agreed (references 5,32) that when measured at the error signal the pilot's remnant can be approximated by white noise added to the signal he is observing over the frequency range of interest, .1 to 30 rad/sec. The white noise was provided by a low frequency noise generator model 301A, made by Elgenco, Inc., of Santa Monica, California. The noise spectrum was found to be flat from below .5 rad/sec to well over 200 rad/sec. In the experiment the noise was injected at the control stick output, c. It was thus necessary to filter the white noise to obtain the desired spectrum at the error. The filter used was a forward gain of 5 with a feedback gain of $1/s$. The symbol CN (colored noise) is used for this filtered noise. The noise, as well as every signal recorded, was passed through a second order filter with two poles at 40 rad/sec in order to avoid aliasing.

The noise spectrum shown on each figure is not flat because it is modified by the loop dynamics, and because it is the result of a single measurement of a random process.

III. RESULTS

In evaluating an identification technique one often processes input-output data from a known model. Results of this type are shown in figures 5, 6, 11 and 17. A perusal of the figures will show that the best results are obtained for this case, a fact which partially explains its wide use (reference 36 for example). Adding noise to the output, adding vehicle dynamics and closing the loop, or using a human being as pilot all reduce the accuracy of identification. Furthermore, these factors and the dynamics identified have different effects on different identification techniques.

Because the main differences between identification techniques

are tied to signal and noise levels, let us briefly review some relevant facts about a pilot in a compensatory tracking task as shown in figure 1:

1) In most reasonable control systems closing the loop reduces the signal level at the error, which is the input to the pilot. Thus the signal to noice level (and the quality of the data) is reduced.

2) Dynamics with high gains over portions of the frequency range cause the pilot to reduce his gain over these frequency ranges. For example, the commonly used $1/s^2$ causes the pilot to reduce his gain below 1 rad/sec (see figure 4), thus degrading the quality of the control stick signal when noise is present.

3) At low frequencies the pilot can control well, hence there is little correlated error power at low frequencies. At high frequencies the pilot cannot control well, and there is thus little correlated control stick power at high frequencies. Only in the middle range, near the crossover frequency, is there high correlated power at both the pilot input and output.

4) Because the pilot cannot control well at high frequencies, common practice (reference 32) is to reduce input power levels above about 3 rad/sec, thus reducing all signal levels at these frequencies.

When noise is added to the output, all the identification techniques do a poorer job, as expected (figures 7 to 10, 12 and 18). The effect is more marked at higher frequencies where both the shelf spectrum and the dynamics reduce signal levels. Closing the loop and adding vehicle dynamics ($1/s$ or $1/s^2$) has a slightly adverse effect on the accuracy of identification (figures 13 and 14, 19 and 20, and Table V). The combination of closing the loop and adding noise has the greatest effect, as seen in figures 15 and 16, 21 and 22, and Table VI. The poorer identification is quite pronounced at low frequencies in figures 21 and 22, where the vehicle dynamics $1/s^2$ greatly reduce the signal level at the error.

For all the runs with known models (figures 5 to 22, 31 and 32) the most accurate results are obtained with techniques 1 and 2 (see Table I), the only techniques to take explicit advantage of the sum of sinusoids forcing function. All the identification techniques do well at crossover, and have an increasing error away from crossover as the signal to noise ratio decreases. Because the parameter trackers assume a fixed model form they do not identify the known models away from

crossover unless the tracker model and the known model forms are the same. The parameter trackers home in at the crossover frequency because this is where the pilot input and output signals both have large signal to noise ratios. The ERC filtering technique gives the poorest performance away from crossover, primarily because of the poor choice of filters.

For the runs with human operators (figures 23 to 30), it is reasonable to assume that technique 1 continues to give the best identification. If such is the case, the previous comments based on identifying known models appear to be valid when identifying the human operator. It should be noted that there is a larger spread between the results obtained by the various identification techniques when identifying a pilot than when identifying pilot models. Thus the human operator is harder to identify than the quasi-linear describing function models (linear describing function plus noise) as used in this experiment.

In working with the two parameter trackers (techniques 5 and 6), one could not help but compare them. Technique 5 measures the pilot only, while technique 6 assumes the crossover model and measures the open loop dynamics. Technique 5 was slightly more stable, converged slightly faster, and could handle more cases than technique 6 (because of the crossover model assumption). The model of technique 6 (equation 12) is more widely accepted by manual control engineers than that of technique 5 (equation 11), but studying figures 23, 25, 27, and 29, one cannot claim either model did a better job of pilot identification than the other.

Adams (references 1,2,3) defines the pilot remnant as the difference between the linear pilot model output and the actual pilot output, while McRuer (reference 32) defines the pilot remnant as that portion of the control stick output which is not correlated with his input. If Adam's model accounts for all the correlated power (of the pilot output with respect to the pilot input), then the two definitions of remnant are the same. Insofar as the parameter tracking model is in error, Adam's definition of remnant will be measured as larger that McRuer's remnant. The concept was experimentally tested as follows:

1) The input data were fed into the Adam's pilot model, giving a pilot model output. The input was the pilot input, or error signal shown in figure 1, and the model used was the one determined for the run corresponding to the data used.

2) The pilot model output was subtracted from the actual pilot output to give a fitting error.

3) The fitting error was analyzed at frequencies ω_k and ω_j (the input sinusoid frequencies and between them).

The experiment was then repeated using the open loop model obtained from technique 6. In both cases the measurements made at frequencies ω_j were identical to those obtained from technique 1 after the latter were corrected for the loop dynamics (that is, converted from in-the-loop levels to injected values outside the loop). Also in both cases, larger powers were measured at frequencies ω_k away from the crossover frequency. Typical results are shown in Table VII. As can be seen, the error of the parameter trackers at the input frequencies away from crossover is in general larger than the measured remnant, thus indicating a fitting error does exist and is larger than the remnant away from crossover.

IV. CONCLUSION

All the identification techniques studied in this experiment provide good measurements in the region of crossover. Among those commonly used, the ones with more restrictions and which are more computationally expensive give the best results away from crossover. For example, techniques 1 and 2 give the best identification and direct measurements of the remnant, but require a sum of sinusoids forcing function and complex data processing. A brief summary is given in Table VIII.

All methods deteriorate when signal levels are low. Closing the loop, controlled dynamics of the form $1/s$ or $1/s^2$, the use of a shelf spectrum, and the pilot's ability to control either very well or poorly, all tend to reduce signal levels at various places in the loop over certain frequency ranges. This leads to the commonly recognized phenomenon of being able to identify the pilot well only in the region of crossover.

BIBLIOGRAPHY

1. Adams, J. J., and H. P. Bergeron, "Measured Variations in the Transfer Function of a Human Pilot in Single Axis Tasks", NASA-TN-D-1952, October,1963.

2. Adams, J. J., "A Simplified Method for Measuring Human Transfer Functions", NASA-TN-D-1782, 1963.

3. Adams, J. J., et al, "Human Transfer Functions in Multi-Axis and Multi-Loop Control Systems", NASA-TN-D-3305, April, 1966.

4. Balakrishnan, A. V., and V. Peterka, "Identification in Automatic Control Systems", Proceedings of the Fourth Congress of the International Federation of Automatic Control, June, 1969.

5. Baron, S., et al, "Application of Optimal Control Theory to Prediction of Human Performance in a Complex Task", Proceedings of the Fifth NASA-University Annual Conference on Manual Control, March, 1969.

6. Baron, S., and D. L. Kleinman, "The Human as an Optimal Controller and Information Processor", NASA-CR-1151, September, 1968.

7. Beare, A. C., and A. Kahn, "Describing Functions for Compensatory Tracking of Sine Waves Plus Noise", Third Annual NASA-University Conference on Manual Control, 1967.

8. Bekey, G. A., et al, "A Study of Model Matching Techniques for the Determination of Parameters in Human Pilot Models", NASA-CR-143, January,1965.

9. Bekey, G. A., "The Human Operator as a Sampled Data System", IRE Transactions, HFE, V. 3, pp 43-51, 1962.

10. Bergland, G. D., "A Guided Tour of the Fast Fourier Transform", IEEE Spectrum, pp 41-52, July, 1969.

11. van den Boom, A. J. W., and J. H. A. M. Melis, "A Comparison of Some Process Parameter Estimating Schemes", Proceedings of the Fourth Congress of the International Federation of Automatic Control, June, 1969.

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

13. Costello, R. G., 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, V. HFE-7, No. 4, pp 174-181, December, 1966

14. Crossman, E.R.F.W., and H. P. Delp, "Application of Gabor's Elementary-Signal Theorem to Estimation of Nonstationary Human Response", presented at the Fifth NASA- University Annual Conference on Manual Control, March 1969.

15. Cuenod, M., and A. P. Sage, "Comparison of Some Methods Used for Process Identification", Automatica, V. 4, No. 4, pp 235-269, May 1968.

16. Eisenberg, L., "Stability of Linear Systems with Transport Lag", IEEE Transactions on Automatic Control, V. AC-11, No. 2, pp 247-254, April 1966.

17. Elkind, J. I., and D. M. Greene, "Measurement of Time-Varying and Non-Linear Dynamic Characteristics of Himan Pilots", ASD Technical Report 61-225, December 1961.

18. Elkind, J. I., et al, "Evaluation of a Technique for Determining Time-Invariant and Time-Variant Dynamic Characteristics of Human Pilots", NASA-TN-D-1897, May 1963.

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

20. Englar, T., "Research Report on Identification of Linear Systems", to appear as a NASA contractors Report.

21. Falb, P. I., and G. Kovatch, "Dynamical Systems Modeling of Human Operators: Preliminary Report", Second NASA-University Annual Conference on Manual Control, NASA-SP-128, February 1966.

22. Ho, Y. C. and B. H. Whelan, "An Approach to the Identification and Control of Linear Dynamic Systems with Unknown Paramaters", IEEE Transactions on Automatic Control, V. AC-8, pp 255-256, July 1963.

23. Hoffman, L. G., et al, "Theoretical and Experimental Research on Parameter Tracking Systems", NASA-CR-452, April 1966.

24. Hsia, T. C., and V. Vimolvanich, "An On-Line Technique for Systems Identification", IEEE Transactions on Automatic Control, V. AC-14, No. 1, pp 92-96, February 1969.

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

26. Kelly, C. R., Manual and Automatic Control, John Wiley and Sons, Inc., 1968.

27. Kelly, C. R., "A Psychological Appraoch to Operator Modeling in Manual Control", Third Annual NASA-University Conference on Manual Control, 1967.

28. Kleinman, D. L., et al, "An Optimal Control Model of Human Response", presented at the Fifth Annual NASA-University Conference on Manual Control, March 1969.

29. Lessing, H. C., and D. F. Crone, "The Use of Integral Transforms in the Estimation of Time Variable Parameters", NASA-TN to be published.

30. Merrit, M. J., "Synthesis and Identification of Mathematical Models for the Discrete Control Behavior of Human Operators", University of Southern California Tech. Report USCEE Report 202, May 1967.

31. Miller, R. A., "A Previous Control Model with One or Two Fast Time Scale Loops", Massachusetts Institute of Technology Master's Thesis, September 1967.

32. Mc Ruer, D., et al, "Human Pilot Dynamics in Compensatory Systems", AFFDL-TR-65-15, July 1965.

33. Mc Ruer, D., and E. S. Krendel, "The Human Operator as a Servo System Element", Journal of the Franklin Institute, V. 267, No.5, May, and No. 6, June 1959.

34. Neal, C. B., "Estimation of the Parameters of Sampled Data Systems by Stochastic Approximation", USC-TR-USCEE-333.

35. Obermeyer, R. W., and F. A. Muckler, "Modern Control System Theory and Human Control Functions", NASA-CR-256, July 1965.

36. Shirley, R. S., "Application of a Modified Fast Fourier Transform to Calculate Human Operator Describing Functions", NASA-TM-X-1762, March 1969.

37. Shirley, R. S., "Some Modern Control Techniques for Human Operator Modeling and Identification", to be published in the proceedings of the Fourth NASA Inter-Center Control Conference, November 4-5, 1969.

38. Smith, F. W., "System Laplace-Transform Estimation from Sampled Data", IEEE Transactions on Automatic Control, V.AC-13, No. 1, pp 37-44, February 1968.

39. Suh, S., "Time Domain Identification of Human Operator Dynamic Response", Grumman Research Dept. Memo RM-392, Grumman Aircraft Engineering Corp., Bethpage, L. I., New York, December, 1967.

40. Suh, S., and G. Zetkov. "Identification of Human Pilot Dynamic Behavior in Single-Axis Tasks Using an On-Line Linear-Model-Matching Technique and Power-Spectral Analysis of the Remnant", Report No. ADR-04-07-65.2, Grumman Aircraft Engineering Corp., Bethpage, L. I., New York, December 1965.

41. Taylor, L. W., "Nonlinear Time-Domain Models of Human Controllers", Fifth Annual NASA-University Conference on Manual Control, March 1969.

42. Taylor, L. W., "A Comparison of Human Response Modelling in the Time and Frequency Domains", Third Annual NASA-University Conference on Manual Control, NASA-SP-144, March 1967.

43. Taylor, L. W., "Discussion of Spectral Human Response Analysis", Second Annual NASA-University Conference on Manual Control, NASA SP-128, February 1966.

44. Tustin, A., "The Nature of the Operator's Response in Manual Control and Its Implications for Controller Design", Journal of the Institute of Electrical Engineers, V. 94, Part IIA, pp 190-202, 1947.

45. Whitbeck, R. F. and F. D. Newell, "Mean Square Estimation of Human Pilot Transfer Functions", Fourth Annual NASA-University Conference on Manual Control, March 1968.

46. Wingrove, R. C., and F. G. Edwards, "A Technique for Identifying Pilot Describing Functions from Routine Flight Test Records", NASA-TN-D-5127, 1969.

47. Wingrove, R. C., and F. G. Edwards, "Measurement of Pilot Describing Functions from Flight Test Data with an Example from Gemini X", IEEE Transactions on Man-Machine Systems, V. MMS-9, No. 3, September 1968.

48. Zeibolz, H., and H. M. Paynter, "Possibilities for a Two-Time Scale Computing System for Control and Simulation of Dynamic Systems", Proc. National Electronics Conference, V. 9, pp 215-223, 1953.

49. Zetkov, G., "Some Spectral Techniques for Identifying Pilot and Other Input-Output Relationships", Report No. ADR-04-07-65.1, Grumman Aircraft Engineering Corp., Bethpage, L. I., New York, September 1965.

50. Zorn, J., "The Evaluation of Fourier Transforms by the Sampling Method", Automatica, V. 4, No. 516, pp 323-335, November 1968.

51. "Second Annual NASA-University Conference on Manual Control", NASA SP-128, 1966.

52. "Third Annual NASA-University Conference on Manual Control", NASA SP-144, 1967.

CLASSIFIED REFERENCE LIST
(numbers refer to bibliography)

General Survey of Identification of Pilots: 4,13,15,18,19,26,42.

Method of Fourier Coefficients: 10,32,33,36,43,44,49.

Other Transform Techniques: 14,29,36,50.

Crossover Pilot Model: 32,33,43.

Time Domain Identification: 12,39,40,41,42,46,47.

Time-Varying Parameters: 17,18.

Parameter Tracking: 1,2,3,11,23,25,40.

System Realization: 20,21,22,37.

Sampled Data Systems: 9,30,34,38.

TABLE I

THE TECHNIQUES FOR MEASURING PILOT FREQUENCY

RESPONSE WHICH WERE STUDIED

no.	type	originator of version used	descriptive remarks	references	symbol
1	frequency analysis	NASA-ERC R. Shirley	uses modified FFT sum of sines input	32,36	×
2	frequency analysis	BBN W. Levison	uses full FFT	34	⊙
3	filtering	NASA-ARC R. Wingrove	orthogonal filters input-output shift	18,46,47	△
4	filtering	NASA-ERC R. Shirley	s-plane filters system realization	22,37	.....
5	parameter tracker	NASA-LRC J. Adams	3-parameter tracker	1,2,3	– – – –
6	parameter tracker	Oakland Univ. G. Jackson	2-parameter tracker cross-over model	25	—·—
7	cross correlation	NASA-ARC R. Wingrove	input-output shift	46,47	⊡

TABLE II

PARAMETERS OF THE INPUT DISTURBANCE USED FOR EVERY RUN

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

Δt = time between interrupts = .02 sec

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

T = period of data-taking = 216 sec

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

TABLE III

s-PLANE POLES USED FOR TECHNIQUE 4*

s = -1.000
s = - .520
s = -1.930
s = - .269
s = -3.700
s = - .722 ± .250j
s = -1.390 ± .450j
s = -2.680 ± .850j

*see Table I

TABLE IV

LIST OF EXPERIMENTAL RUNS MADE

dynamics $Y_p(s)$	vehicle dynamics $Y_c(s)$	run no.	noise *	OL or CL	see figure	data on figure	techniques used **
(1/s+1) and $(1/s+1)^2$	none	1	NN	OL	2/a	5,6	1,4
	none	2	WN	OL	2/a	7,8	1,4
	none	3	CN	OL	2/a	9,10	1,4
see 1 below	none	4	NN	OL	2/a	11	1,4,5
	none	5	CN	OL	2/a	12	1,4,5
	1/s	6	NN	CL	2/b	13,14	1,4,5,6
	1/s	7	CN	CL	2/b	15,16	1,4,5,6
see 2 below	none	8	NN	OL	2/a	17	1,3,4,5,7
	none	9	CN	OL	2/a	18	1,3,4,5,7
	$1/s^2$	10	NN	CL	2/b	19,20	1,3,4,5,6,7
	$1/s^2$	11	CN	CL	2/b	21,22	1,3,4,5,6,7
subject no. 1	1/s	12	-	CL	1	23,24	1,3,4,5,6,7
	$1/s^2$	13	-	CL	1	25,26	1,3,4,5,6,7
subject no. 2	1/s	14	-	CL	1	27,28	1,4,5,6
	$1/s^2$	15	-	CL	1	29,30	1,4,5,6
$(1/s+1)^2$	none	16	NN	OL	2/a	31,32	1,2,4
subject no. 1	1/s	17	-	CL	1	33,34	1,2,4
	$1/s^2$	18	-	CL	1	35,36	1,2,4

1: $$5.0\left(\frac{8.7 - s}{8.7 + s}\right) = Y_p(s)$$

2: $$.75\left(\frac{8 - s}{8 + s}\right)\left(\frac{4.4s + 1}{.04s + 1}\right)\left(\frac{1}{.12s + 1}\right) = Y_p(s)$$

* NN = no noise added to output
WN = white noise added to output
CN = colored noise added to output

** see Table I

TABLE V

MEASUREMENT ERRORS, RUN 10, NO NOISE

Error of Measured Amplitude Ratio (%)							
freq. no.	freq. rad/s	FT NASA-ERC (no. 1)*	PT NASA-LRC (no. 5)	PT Oakland (no. 6)	Filt. NASA-ERC (no. 4)	Filt. NASA-ARC (no. 3)	Cr.-Cor. NASA-ARC (no. 7)
1	.1164	5	142	56	167	13	17
2	.1745	2	126	41	147	7	12
3	.2909	7	100	24	100	6	6
4	.4363	8	61	14	61	1	7
5	.5818	2	48	10	42	3	1
6	.8727	1	28	6	23	3	2
7	1.309	0	16	3	14	3	4
8	1.745	1	9	1	10	2	1
9	2.618	0	3	2	10	4	4
10	4.363	5	23	11	5	4	1
11	6.545	4	43	27	26	5	5
12	8.727	3	54	49	53	1	5
13	15.71	7	75	145	83	9	9
14	26.18	8	77	358	94	18	9

Error of Measured Phase (degrees)							
freq. no.	freq. rad/s	FT NASA-ERC (no. 1)	PT NASA-LRC (no. 5)	PT Oakland (no. 6)	Filt. NASA-ERC (no. 4)	Filt. NASA-ARC (no. 3)	Cr.-Cor. NASA-ARC (no. 7)
1	.1164	-8	-15	63	-16	4	-5
2	.1745	0	-20	52	-21	6	-5
3	.2909	-7	-24	37	-26	5	-6
4	.4363	4	-24	26	-26	4	-5
5	.5818	-1	-22	18	-24	3	-5
6	.8727	0	-17	10	-18	1	-3
7	1.309	-1	-12	4	-10	0	-2
8	1.745	-2	-7	1	-5	-2	-3
9	2.618	-2	-1	-1	1	0	0
10	4.363	-2	11	3	-2	-5	-5
11	6.545	-4	25	16	-10	1	-1
12	8.727	-1	40	31	-9	-5	-3
13	15.71	0	76	70	14	-4	-1
14	26.18	2	106	102	43	9	-13

* see Table I

TABLE VI

MEASUREMENT ERRORS FOR RUN 11 DUE TO NOISE ONLY

(corrected for errors in run 10)

Error of Measured Amplitude Ratio (%)							
freq. no.	freq. rad/s	FT NASA-ERC (no. 1)*	PT NASA-LRC (no. 5)*	PT Oakland (no. 6)*	Filt. NASA-ERC (no. 4)*	Filt. NASA-ARC (no. 3)*	Cr.-Cor. NASA-ARC (no. 7)*
1	.1164	143	3	0	30	48	45
2	.1745	66	8	0	18	34	40
3	.2909	18	5	0	17	25	25
4	.4363	2	3	0	16	17	12
5	.5818	7	3	0	9	10	9
6	.8727	3	0	0	7	6	4
7	1.309	1	0	0	3	5	1
8	1.745	1	1	0	3	4	2
9	2.618	2	0	0	2	3	1
10	4.363	1	0	0	0	4	6
11	6.545	1	1	0	2	0	5
12	8.727	2	1	1	0	1	15
13	15.71	1	3	5	0	3	3
14	26.18	3	0	1	0	4	91
Error of Measured Phase (degrees)							
freq. no.	freq. rad/s	FT NASA-ERC (no. 1)*	PT NASA-LRC (no. 5)*	PT Oakland (no. 6)*	Filt. NASA-ERC (no. 4)*	Filt. NASA-ARC (no. 3)*	Cr.-Cor. NASA-ARC (no. 7)*
1	.1164	17	0	0	-1	-11	-6
2	.1745	14	0	0	-2	-13	-8
3	.2909	4	0	0	-2	-7	-1
4	.4363	-5	0	0	-3	-7	-2
5	.5818	-3	0	0	-2	-5	0
6	.8727	-10	0	0	-2	-3	0
7	1.309	1	0	0	-2	-2	0
8	1.745	0	0	0	-1	1	4
9	2.618	0	0	0	0	0	0
10	4.363	0	0	0	2	7	12
11	6.545	0	0	0	3	2	3
12	8.727	-1	0	0	6	11	14
13	15.71	3	0	0	10	14	15
14	26.18	-1	0	0	16	7	20

*see Table I

TABLE VII

REMNANT AND FITTING ERRORS OF PARAMETER TRACKERS

	Run 12 (pilot controlling 1/s)** Measured Power Per Measurement Window			
freq. rad/sec	Remnant at c technique 1*	Fitting Error technique 5*	Remnant at m technique 1*	Fitting Error technique 6*
.116	.002	.036	.069	.026
.175	.003	.149	.080	.070
.291	.001	.404	.016	.070
.436	.004	.070	.027	1.14
.582	.008	.067	.027	.630
.873	.011	.264	.021	1.69
1.31	.030	.768	.015	2.35
1.75	.094	2.59	.026	1.64
2.62	.131	18.5	.023	2.03
4.36	1.49	1.07	.056	.094
6.55	1.78	1.02	.055	1.70
8.73	.921	.341	.015	.941
15.7	.172	.345	.001	.181
26.2	.021	.054	0	.024
	Run 13 (pilot controlling $1/s^2$)*** Measured Power Per Measurement Window			
freq. rad/sec	Remnant at c technique 1*	Fitting Error technique 5*	Remnant at m technique 1*	Fitting Error technique 6*
.116	.006	.127	.210	.380
.175	.003	.420	.340	1.34
.291	.005	1.30	.540	2.84
.436	.001	2.16	.150	.071
.582	.034	.260	.130	1.75
.873	.410	.540	.520	2.44
1.31	1.04	.320	.400	.990
1.75	4.27	4.14	.270	5.96
2.62	26.2	13.3	.470	14.3
4.36	51.1	9.1	.230	2.43
6.55	29.2	37.3	.040	.360
8.73	2.40	38.9	.001	.190
15.7	2.90	11.8	.0002	.032
26.2	.100	1.66	0	.004

*see Table I
**see figures 23 and 24
***see figures 25 and 26

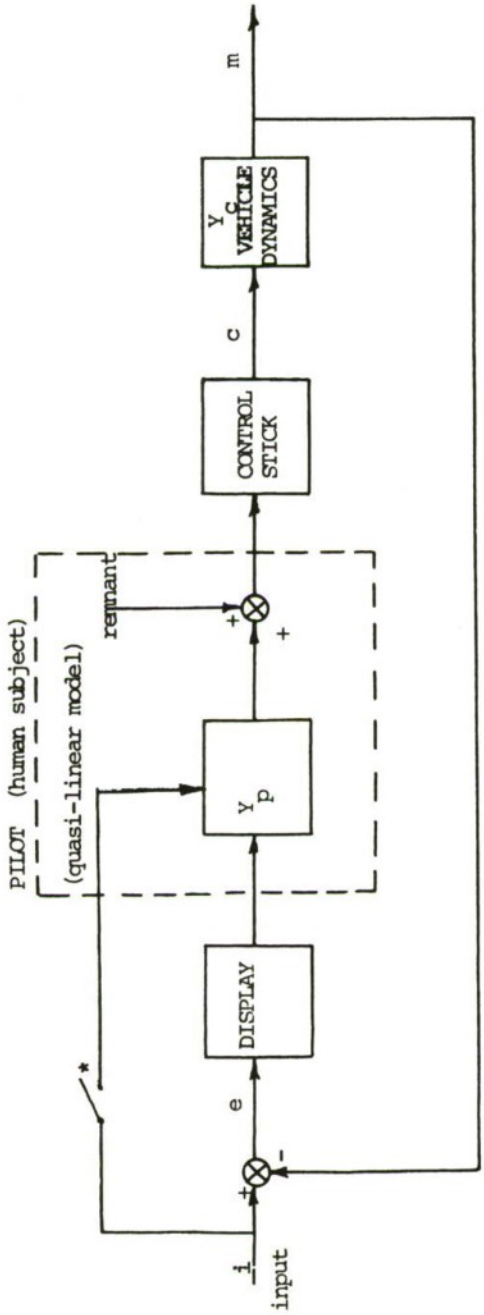
TABLE VIII

CAPABILITIES OF THE IDENTIFICATION TECHNIQUES

Technique Number*	Remnant Def. (1 or 2)**	Random Input Possible?	Accuracy Away From ω_{co}	Gives Model of Pilot?	Other Comments
1	2	no	good	no	results 5 seconds after run
2	2	no	good	no	best measure of remnant
3	1	yes	good	yes	uses time shift to reduce noise effects
4	1	yes	poor	yes	uses B.L. Ho. algorithm (modified)
5	1	yes	OK	yes	easy to implement, fairly stable measures pilot only
6	1	yes	OK	yes	easy to implement specialized to crossover model
7	1	yes	good	no	uses time shift to reduce noise effects

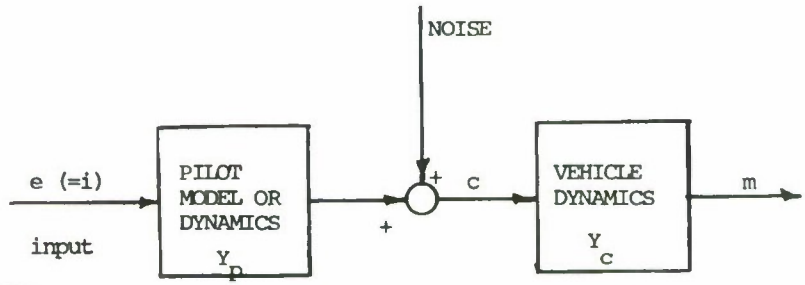
*see Table II

**Definitions of remnant:
1) the difference between the actual pilot output and the pilot model output to the same input
2) the portion of the pilot output which is uncorrelated with his input



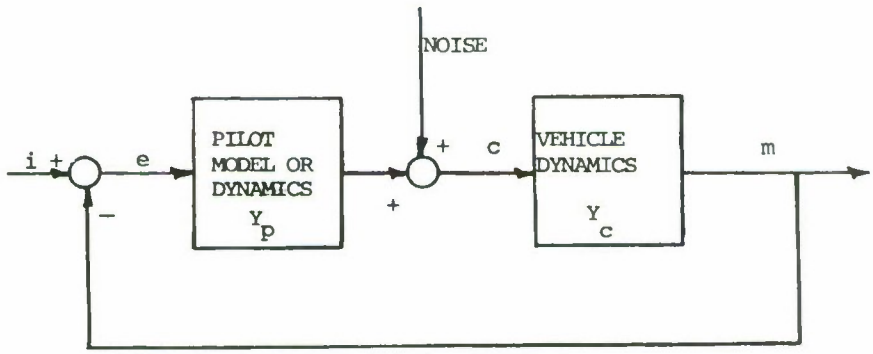
THE PILOT IN A COMPENSATORY OR PURSUIT TRACKING TASK

figure 1



2/a

OPEN LOOP



2/b

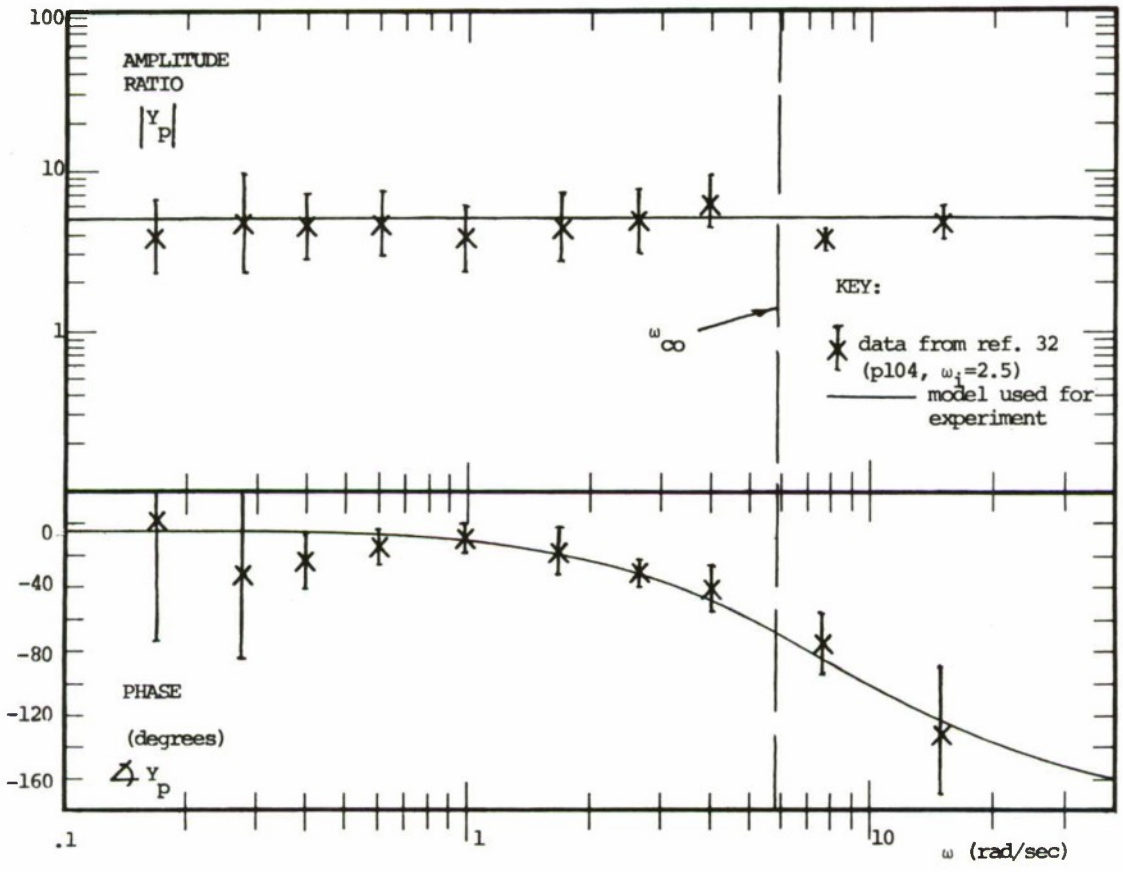
CLOSED LOOP

notes: Y_p and Y_c are fixed analog filters here

measurements are taken at e, c and m

BLOCK DIAGRAM FOR RUNS 1 TO 11 AND 16

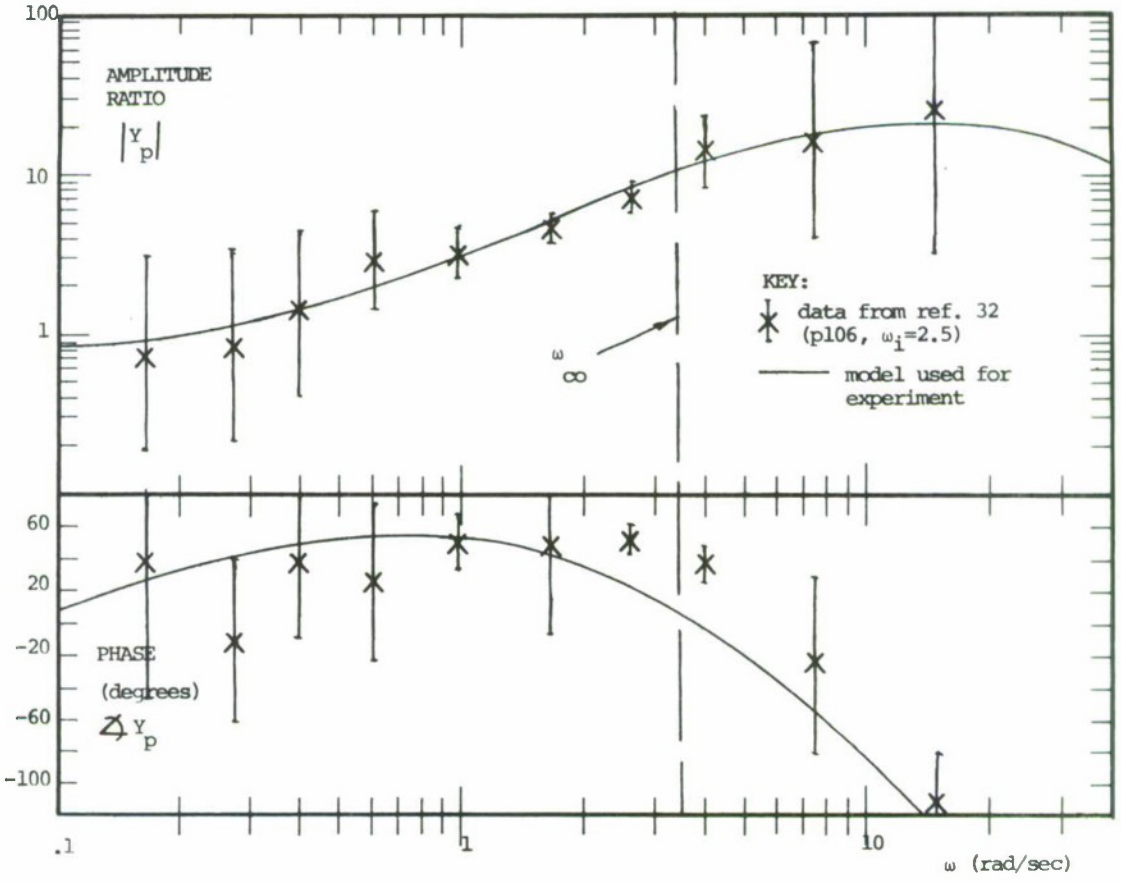
figure 2



COMPARISON OF PILOT MODEL USED FOR EXPERIMENT WITH

DATA FROM TR-65-15 (ref. 32) FOR Y_c = 1/s

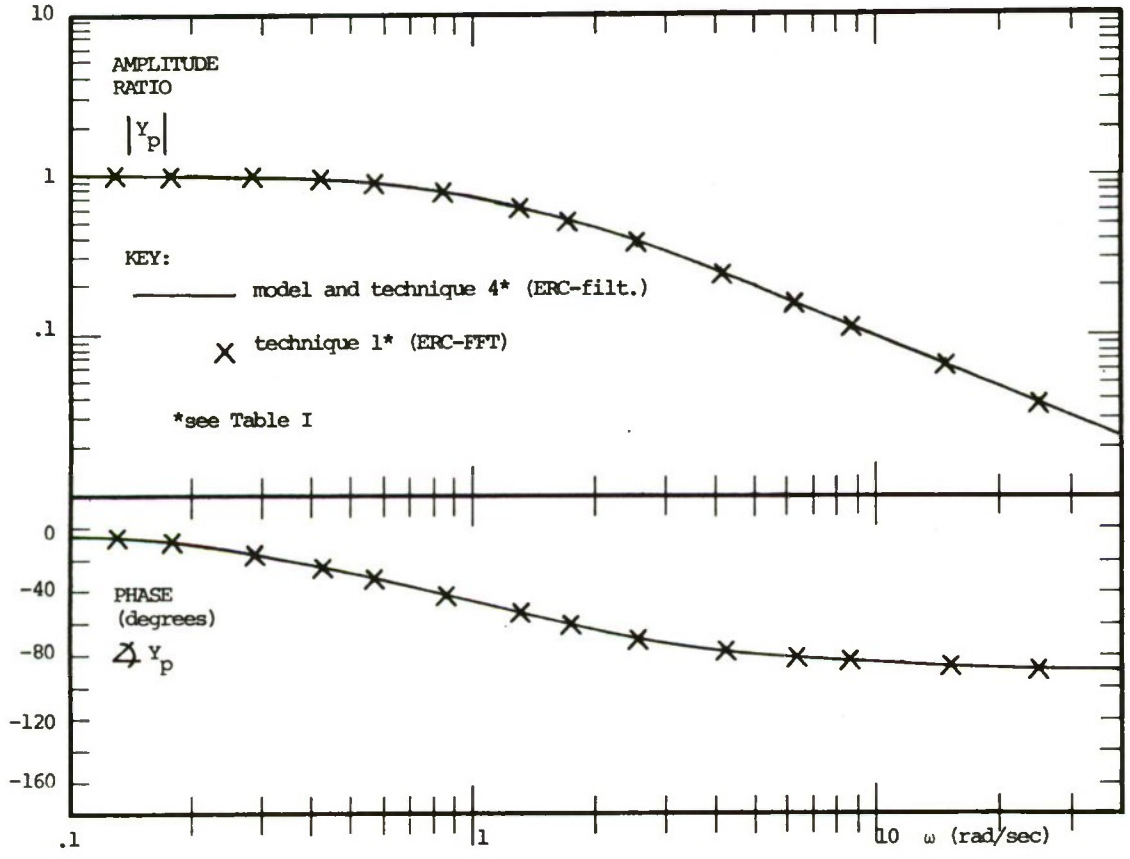
figure 3



COMPARISON OF PILOT MODEL USED FOR EXPERIMENT WITH

DATA FROM TR-65-15 (ref. 32) FOR $Y_c = 1/s^2$

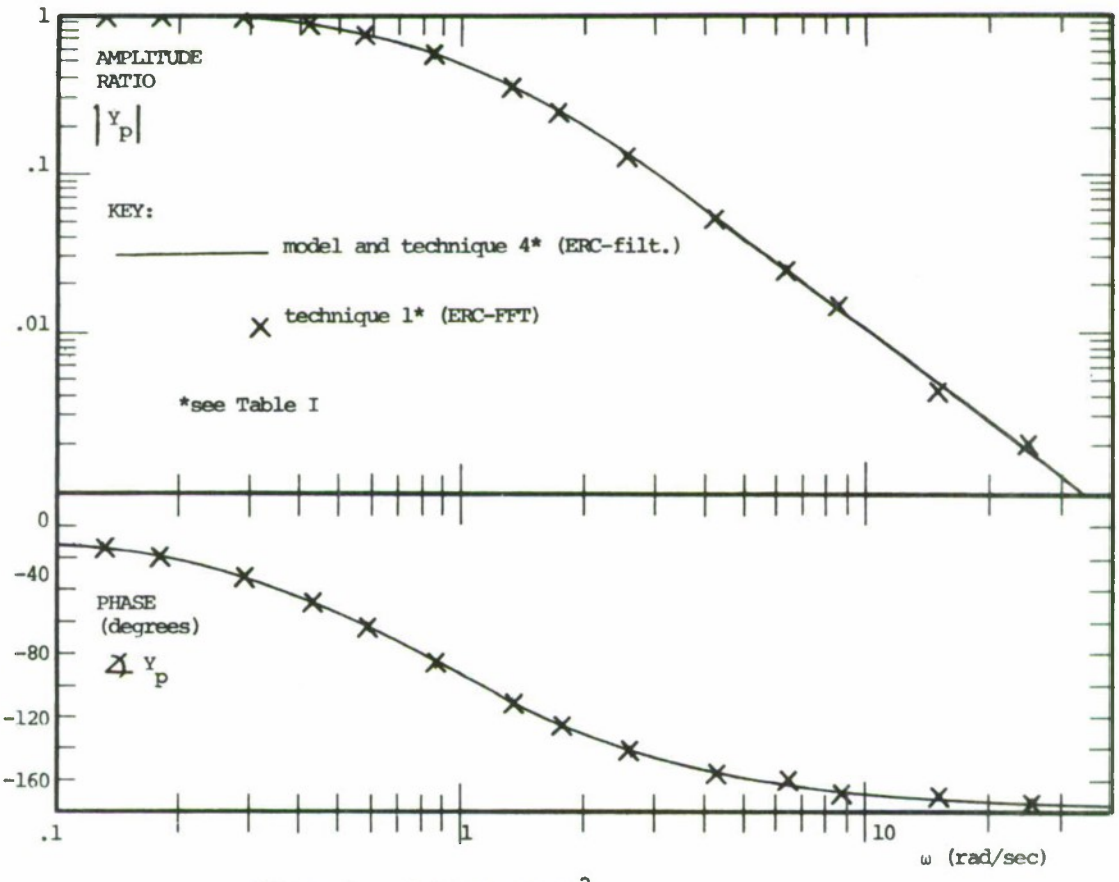
figure 4



RUN NO. 1: $Y_p(s) = (1/s+1)$;

OPEN LOOP NO NOISE

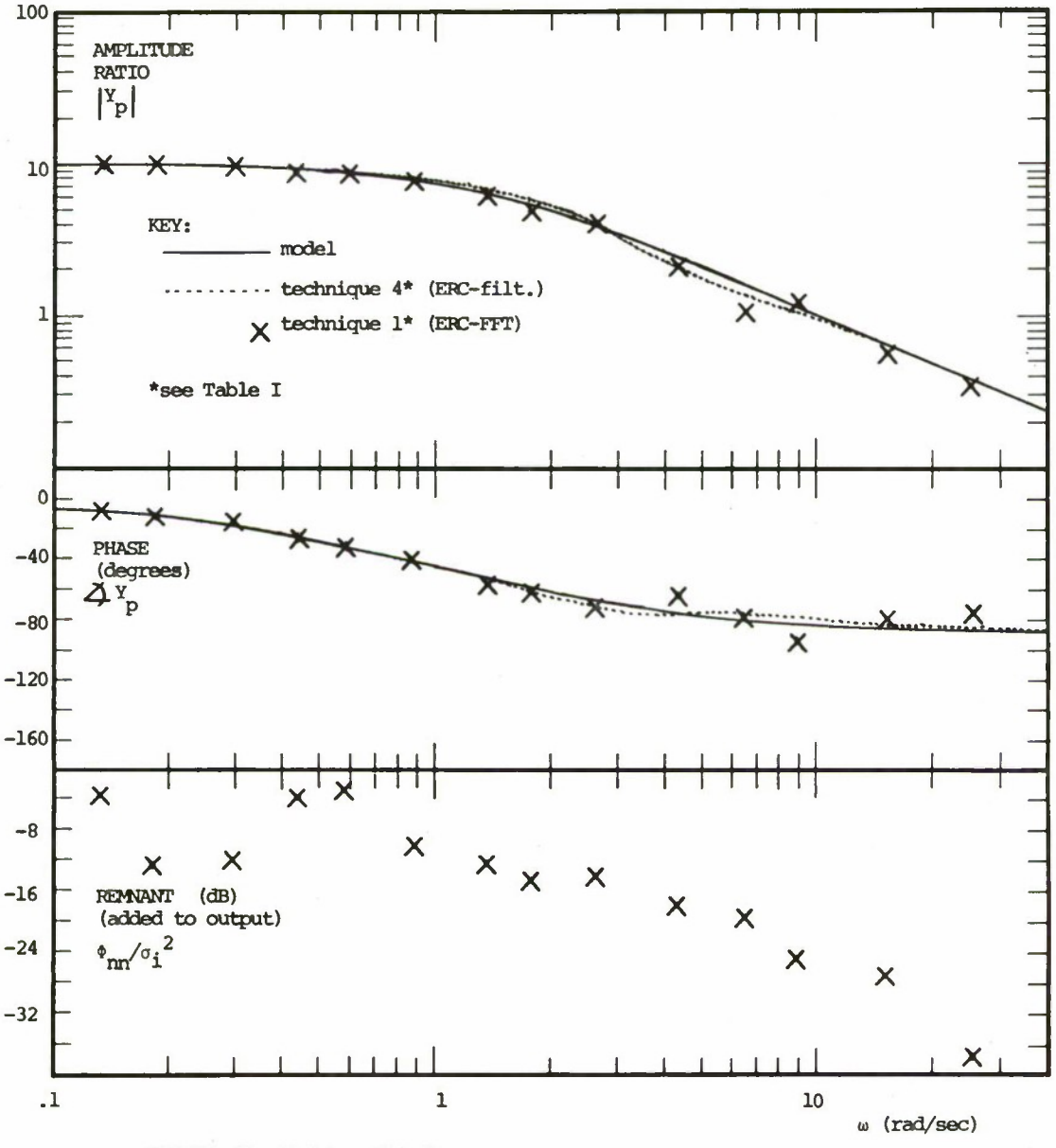
figure 5



RUN NO. 1: $Y_p(s) = (1/s+1)^2$;

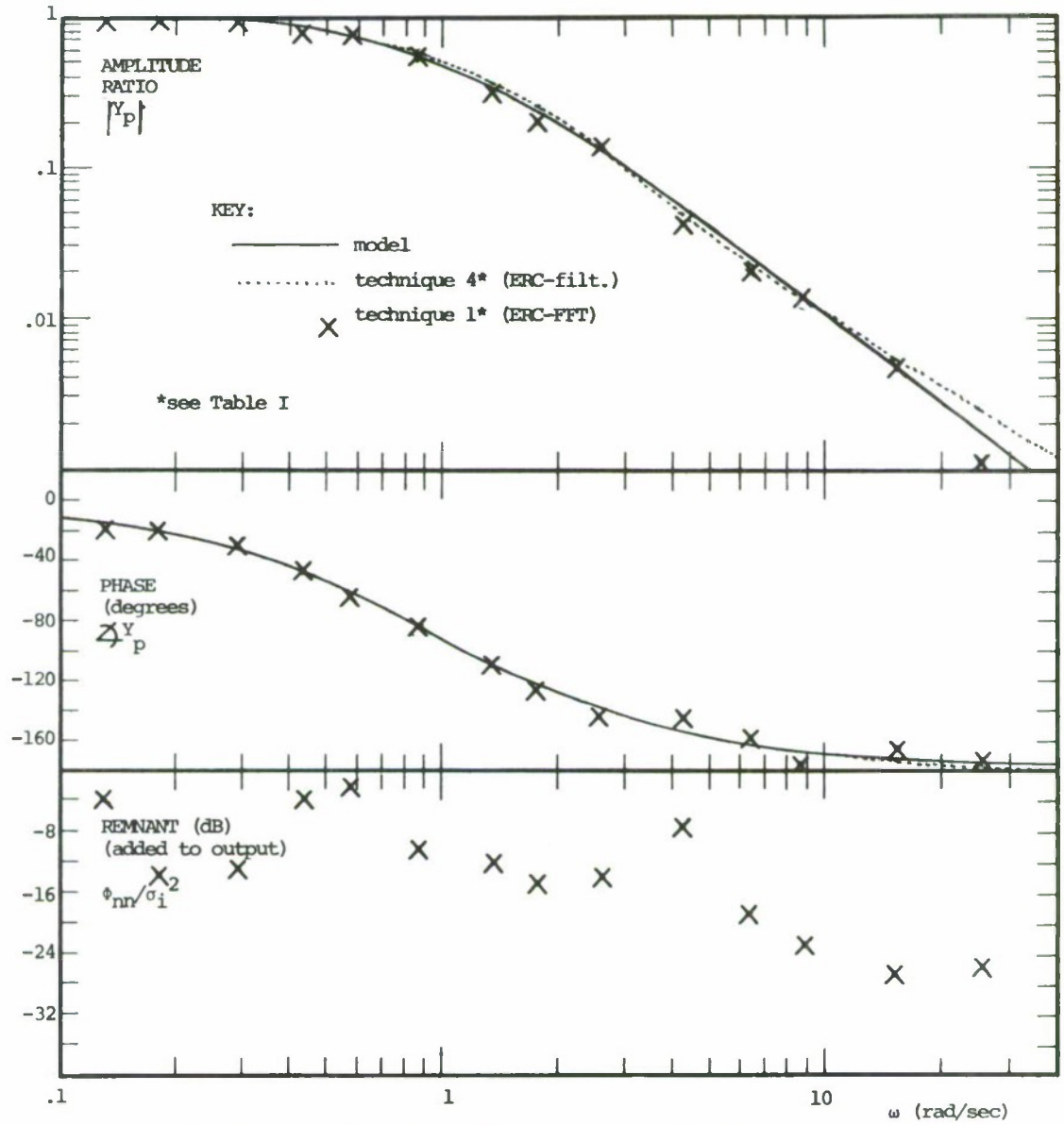
OPEN LOOP NO NOISE

figure 6



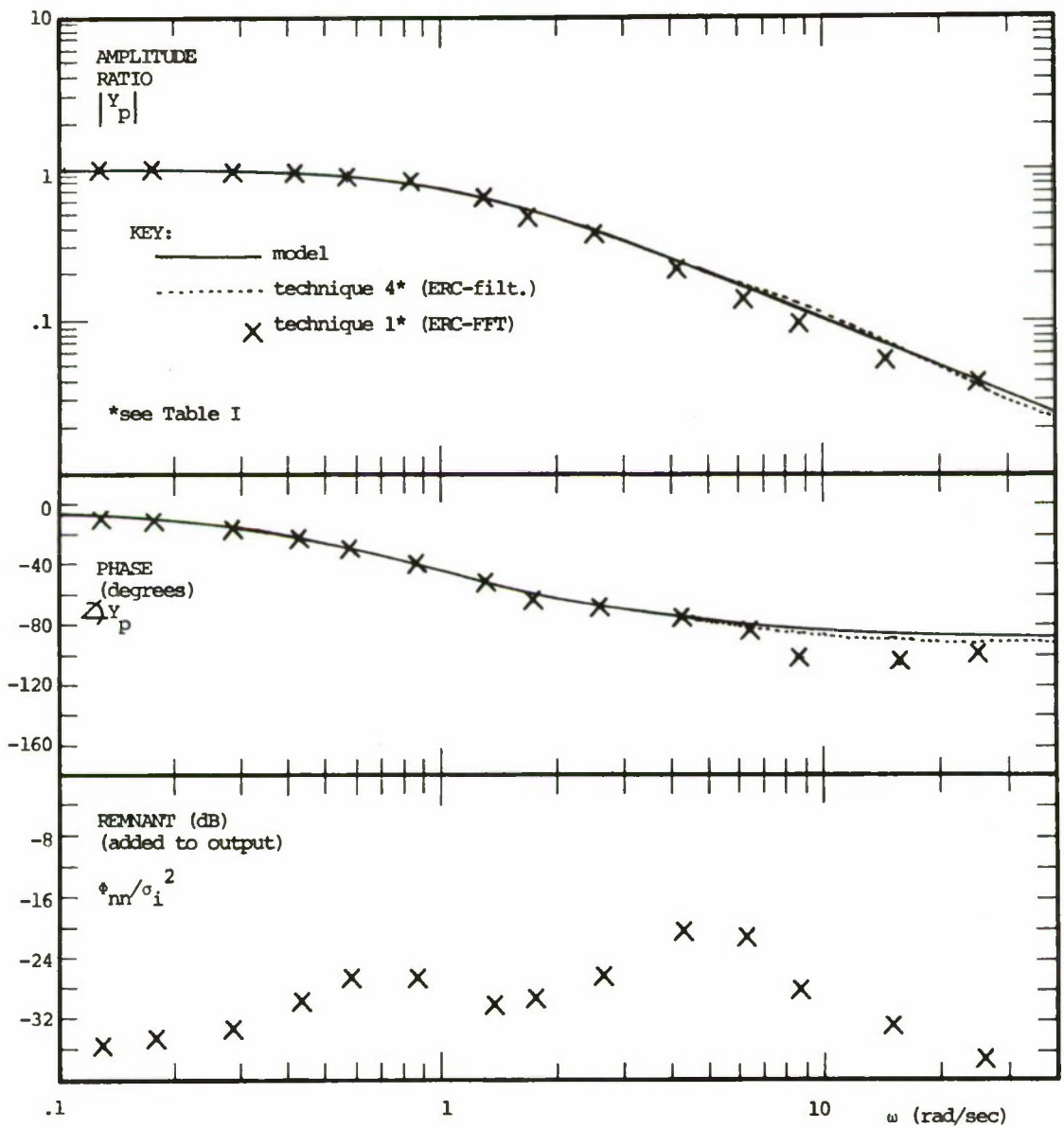
RUN NO. 2 $Y_p(s) = (1/s+1)$; OPEN LOOP

figure 7



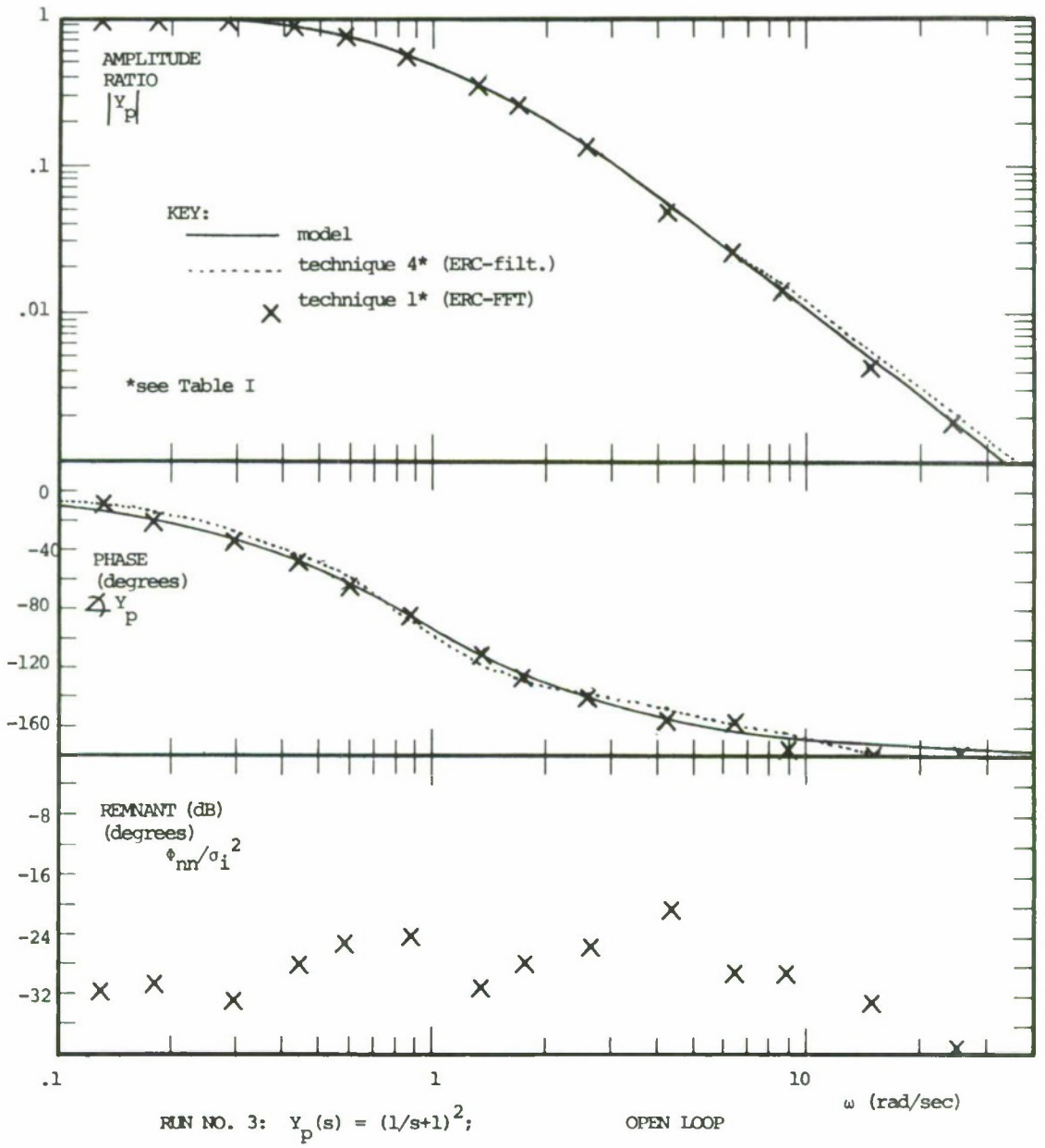
RUN NO. 2: $Y_p(s) = (1/s+1)^2$; OPEN LOOP

figure 8



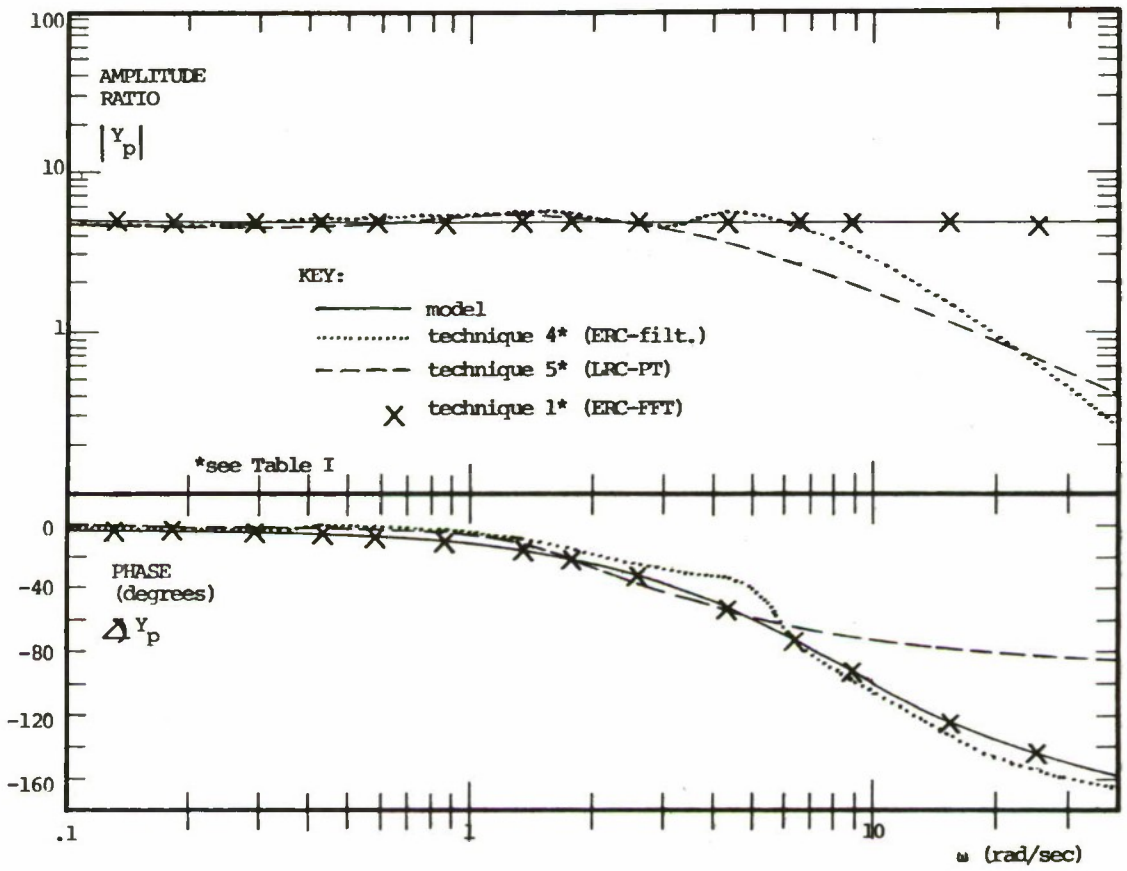
RUN NO. 3: $Y_p(s) = (1/s+1)$; OPEN LOOP

figure 9



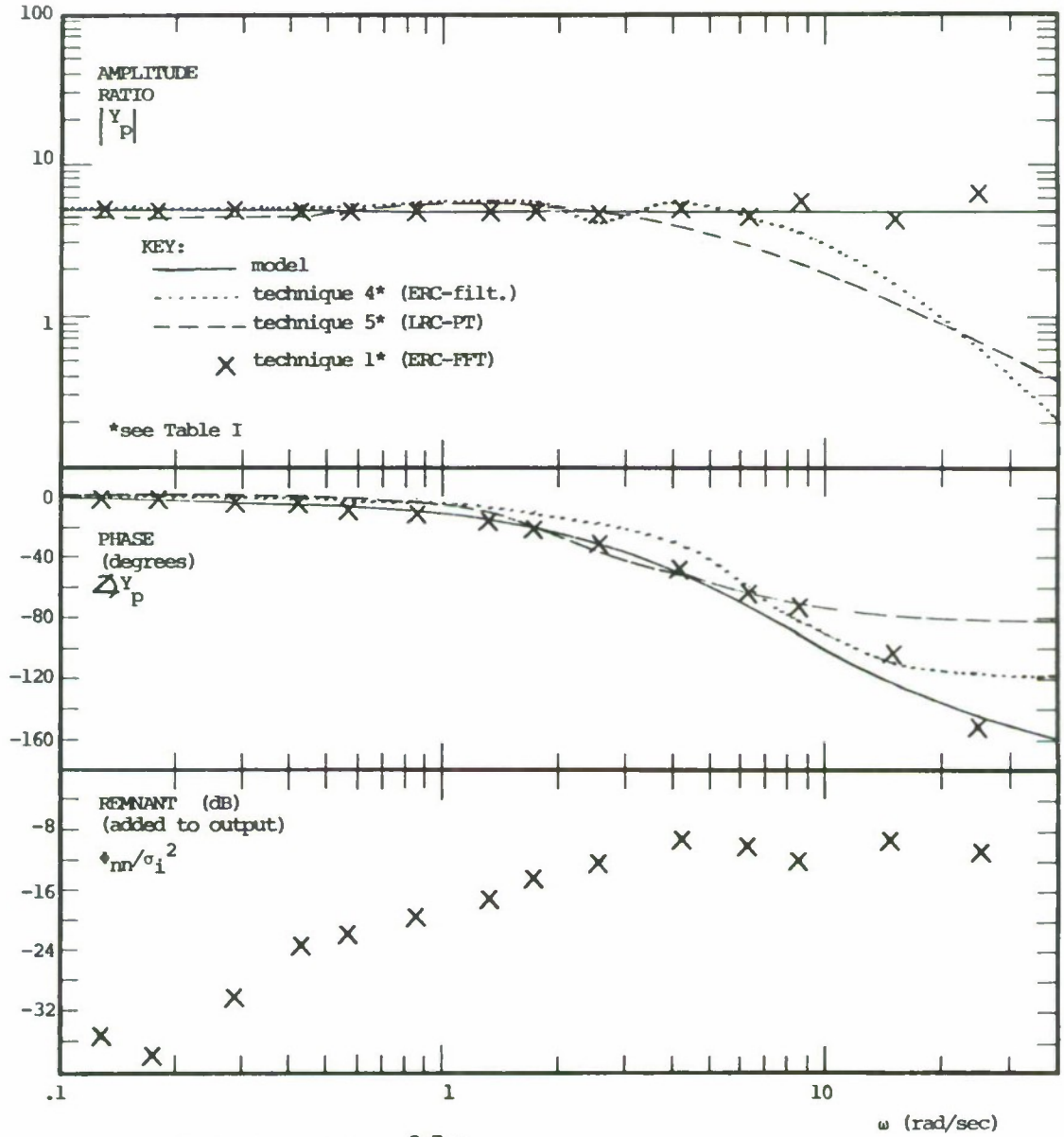
RUN NO. 3: $Y_p(s) = (1/s+1)^2$; OPEN LOOP

figure 10



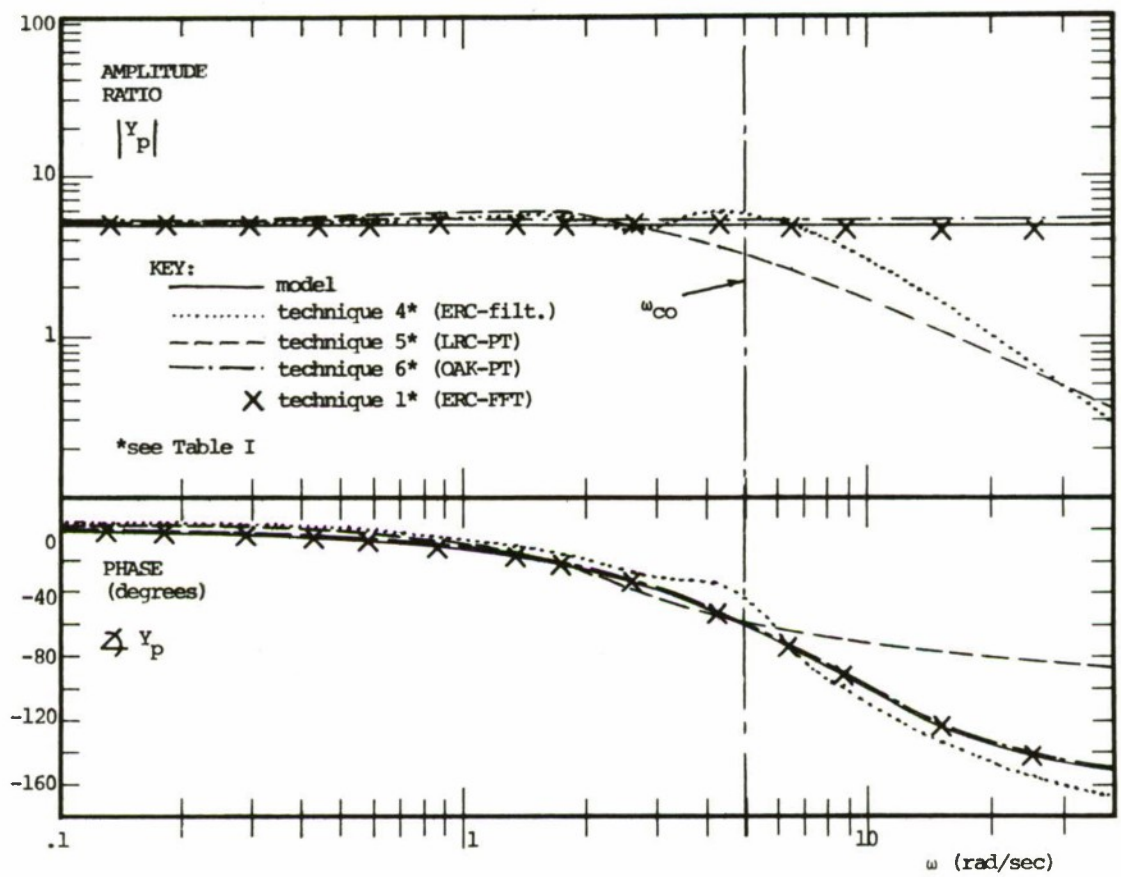
RUN NO. 4: $Y_p(s) = 5\,\frac{8.7-s}{8.7+s}$; OPEN LOOP; NO NOISE

figure 11



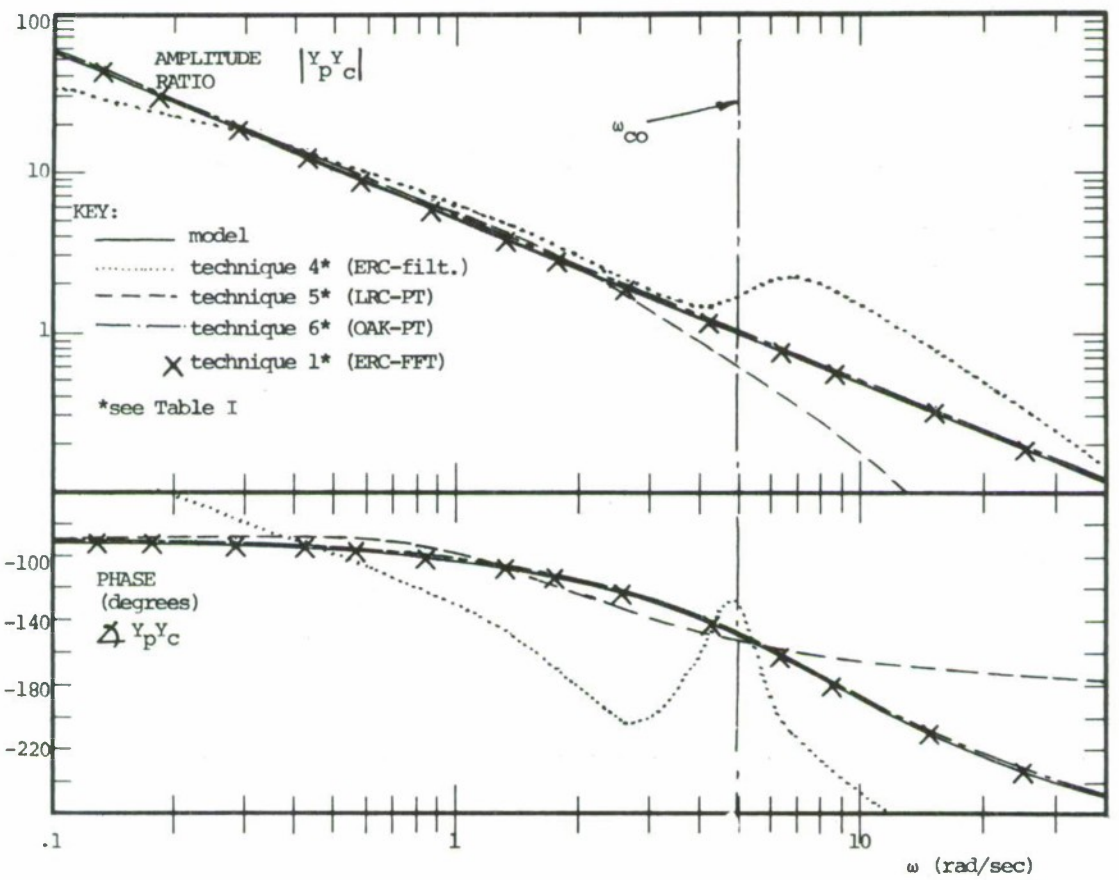
RUN NO. 5: $Y_p(s) = 5\,\frac{8.7-s}{8.7+s}$; OPEN LOOP

figure 12



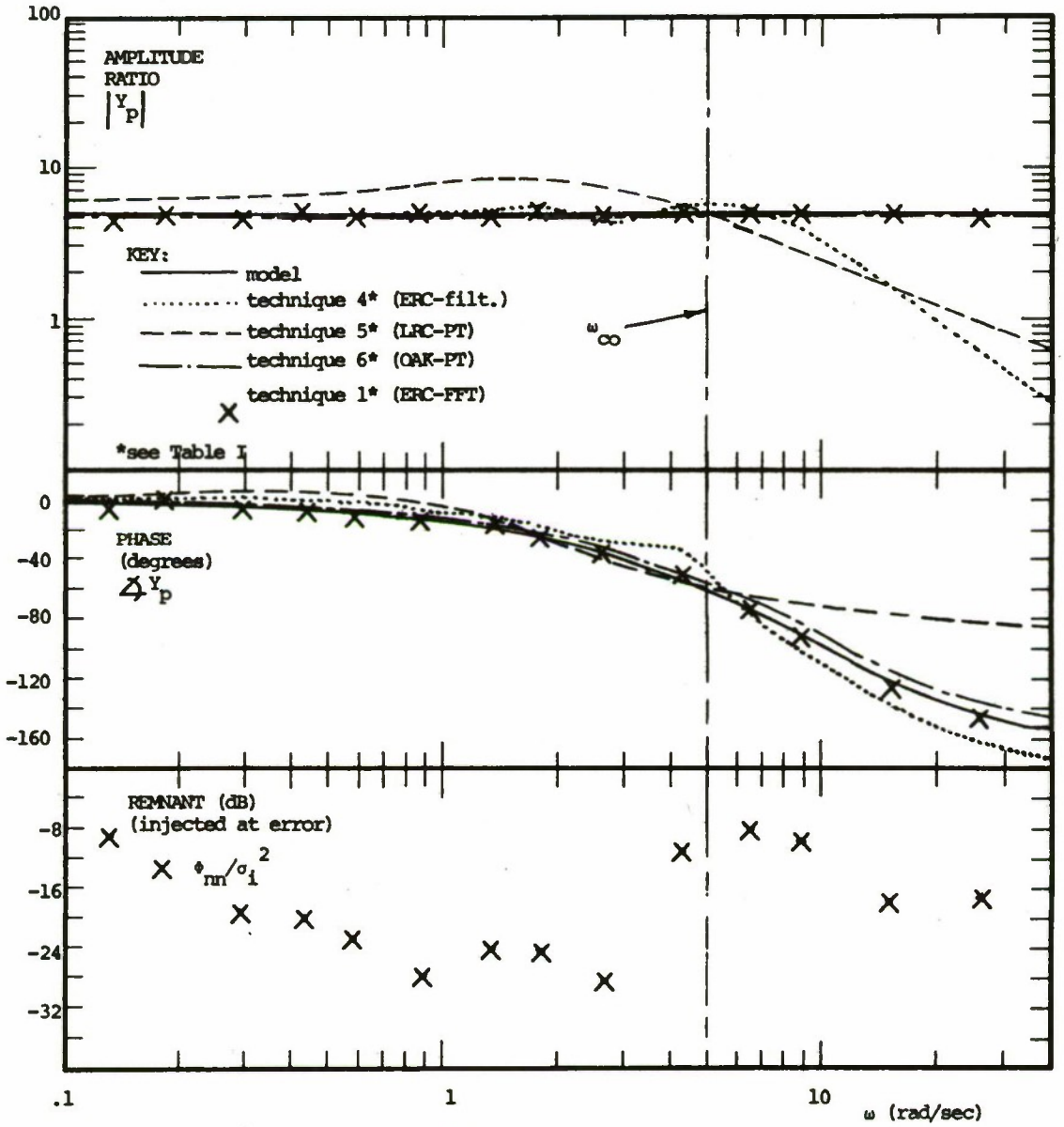
RUN NO. 6: $Y_p(s) = 5\,\frac{8.7-s}{8.7+s}$; $Y_c(s) = 1/s$; CLOSED LOOP; NO NOISE

figure 13



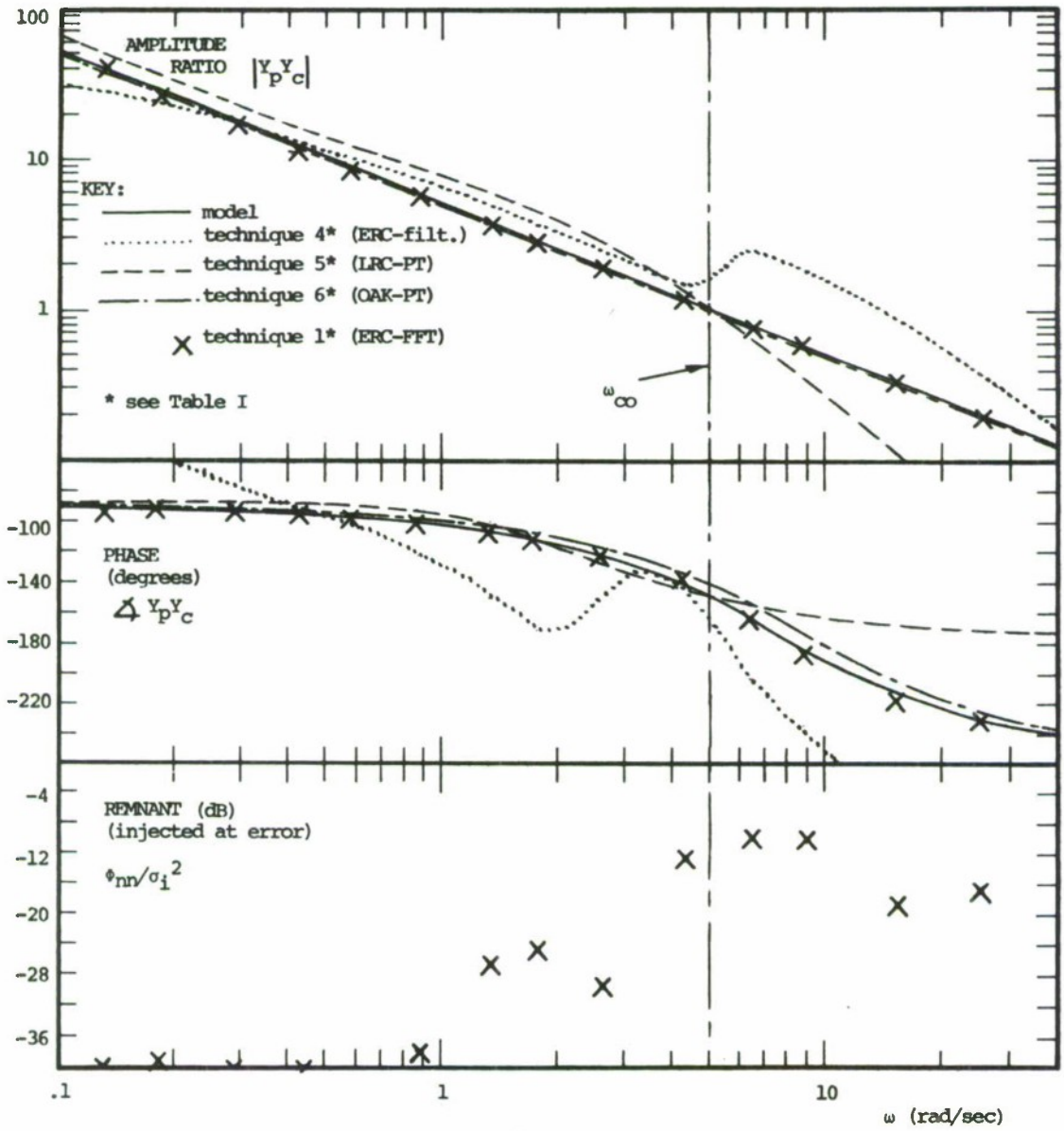
RUN NO. 6: $Y_p(s) = 5\,\frac{8.7-s}{8.7+s}$; $Y_c(s) = 1/s$; CLOSED LOOP; NO NOISE

figure 14



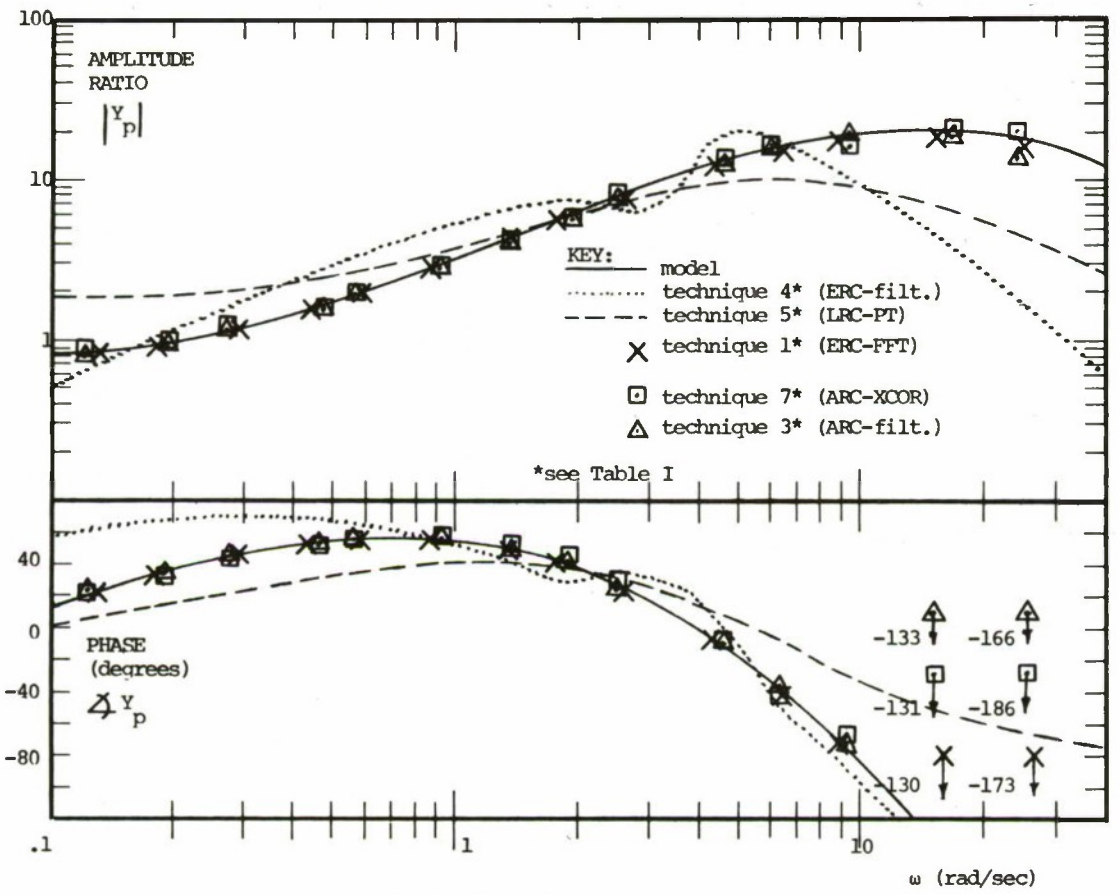
RUN NO. 7: $Y_p(s) = 5\,\frac{8.7-s}{8.7+s}$; $Y_c(s) = 1/s$; CLOSED LOOP

figure 15



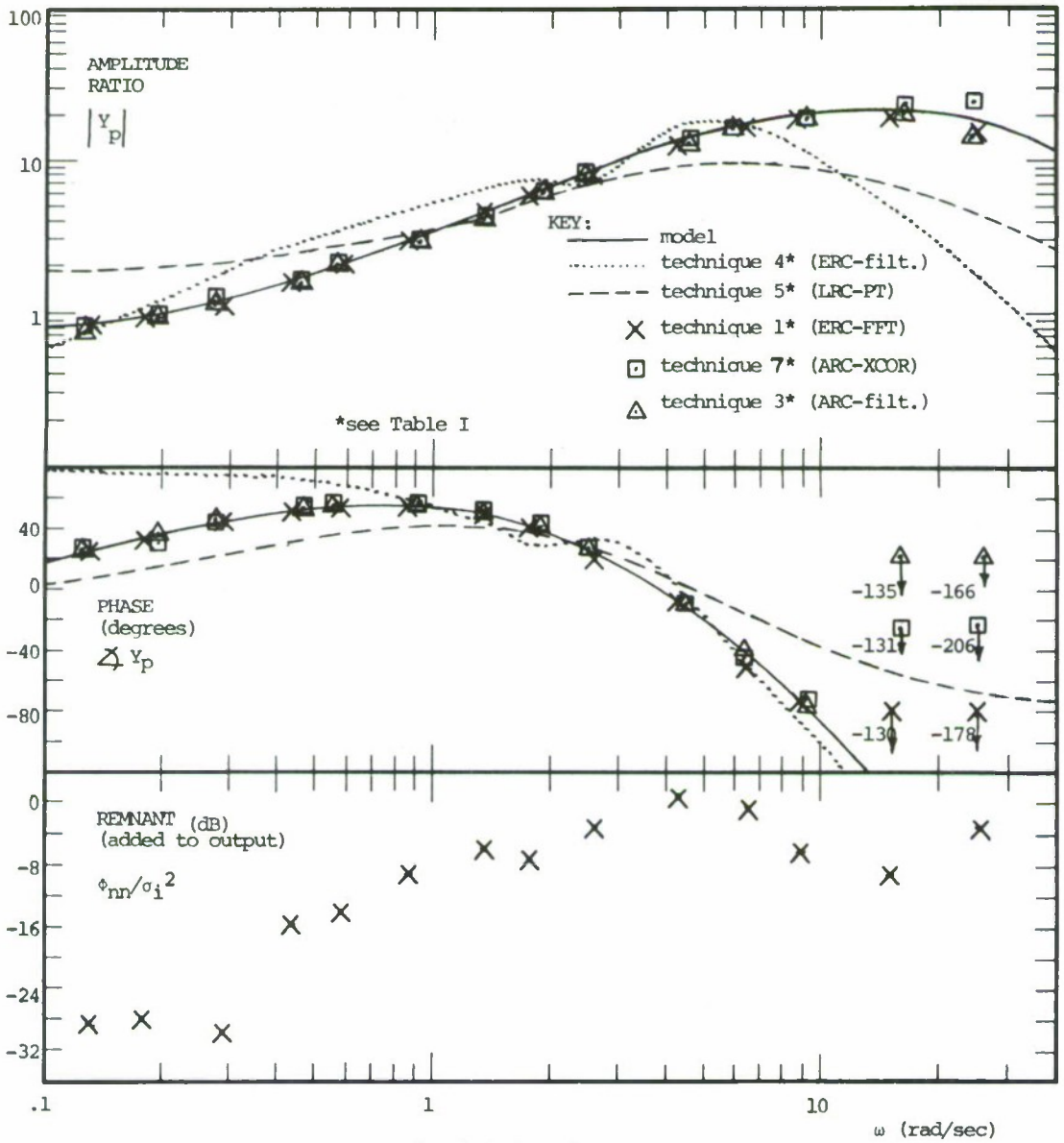
RUN NO. 7: $Y_p(s) = 5\ \frac{8.7-s}{8.7+s}$; $Y_c(s) = 1/s$ CLOSED LOOP

figure 16



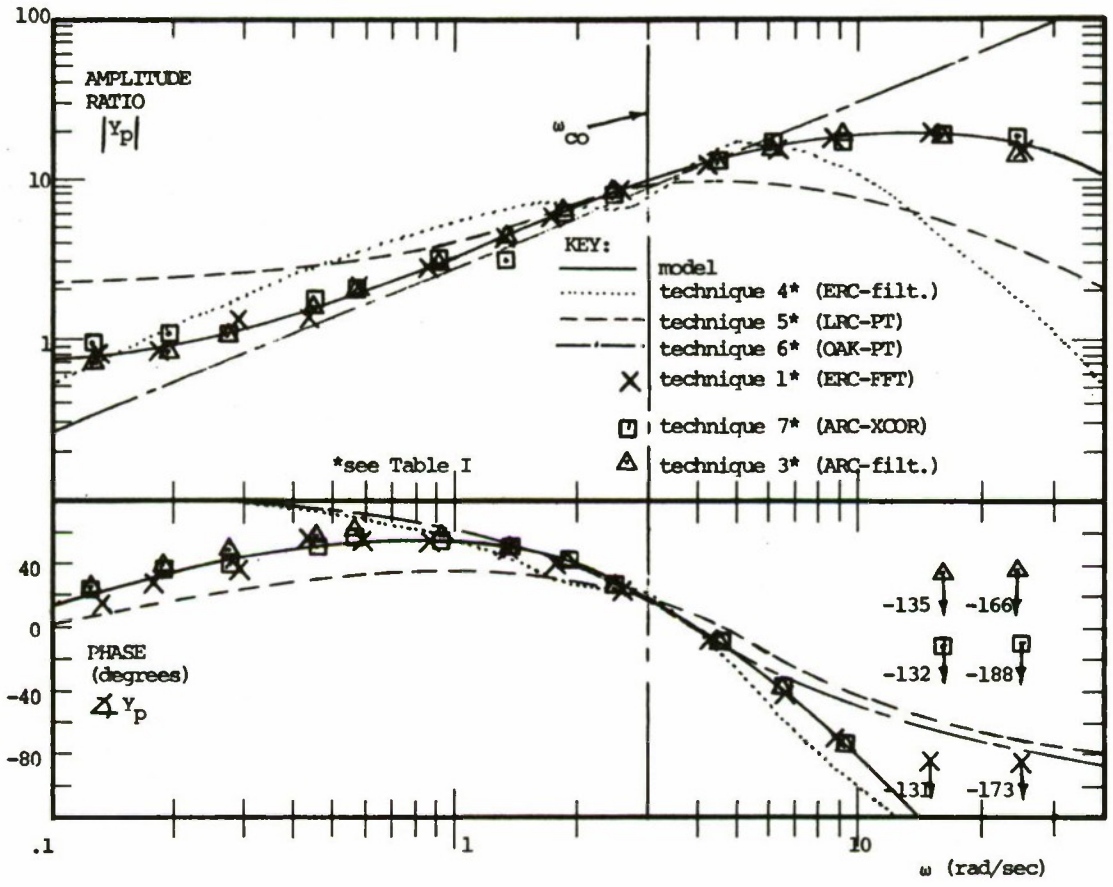
RUN NO. 8: $Y_p(s) = .75 \frac{8-s}{8+s} \frac{4.4s+1}{.04s+1} \frac{1}{.12s+1}$; OPEN LOOP; NO NOISE

figure 17



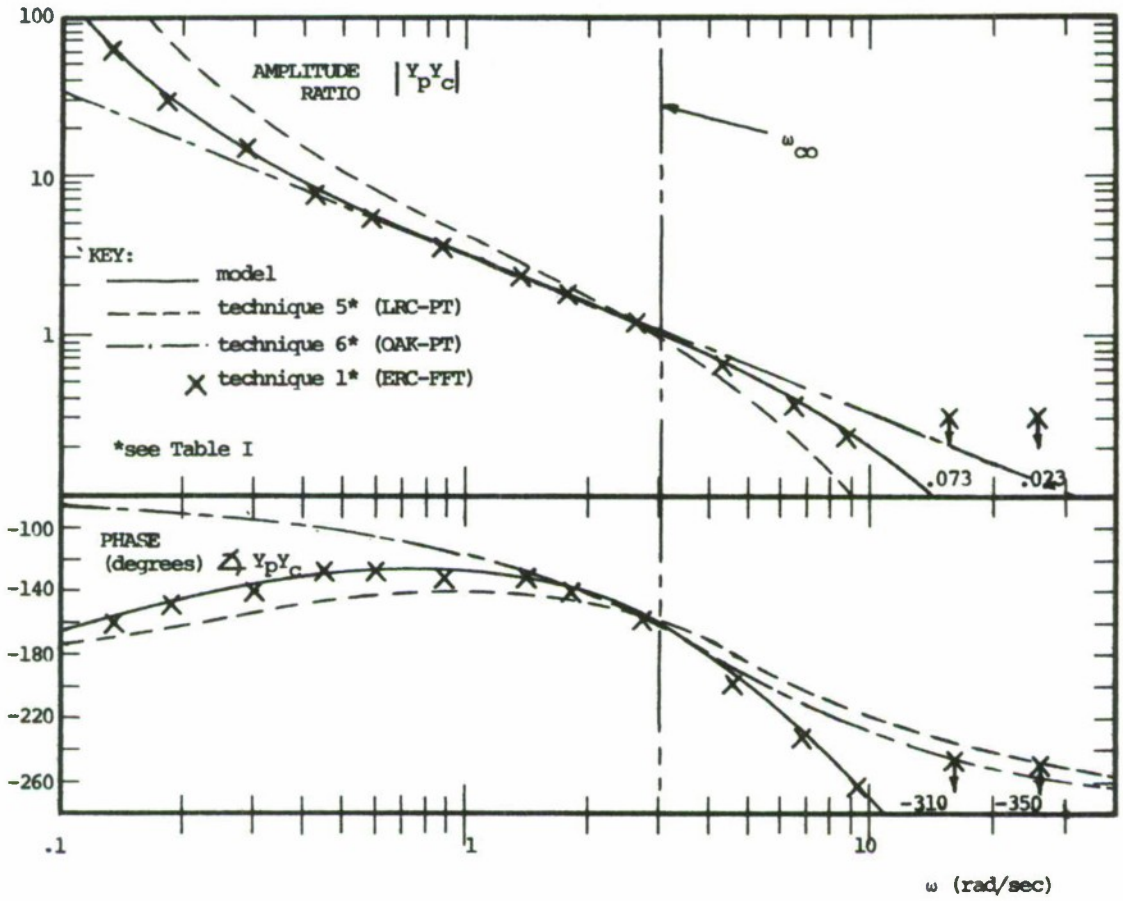
RUN NO. 9: $Y_p(s) = .75 \frac{8-s}{8+s} \frac{4.4s+1}{.04s+1} \frac{1}{.12s+1}$; OPEN LOOP

figure 18



RUN NO. 10: $Y_p(s) = .75 \frac{8-s}{8+s} \frac{4.4s+1}{.04s+1} \frac{1}{.12s+1}$; $Y_c(s) = 1/s^2$ CLOSED LOOP; NO NOISE

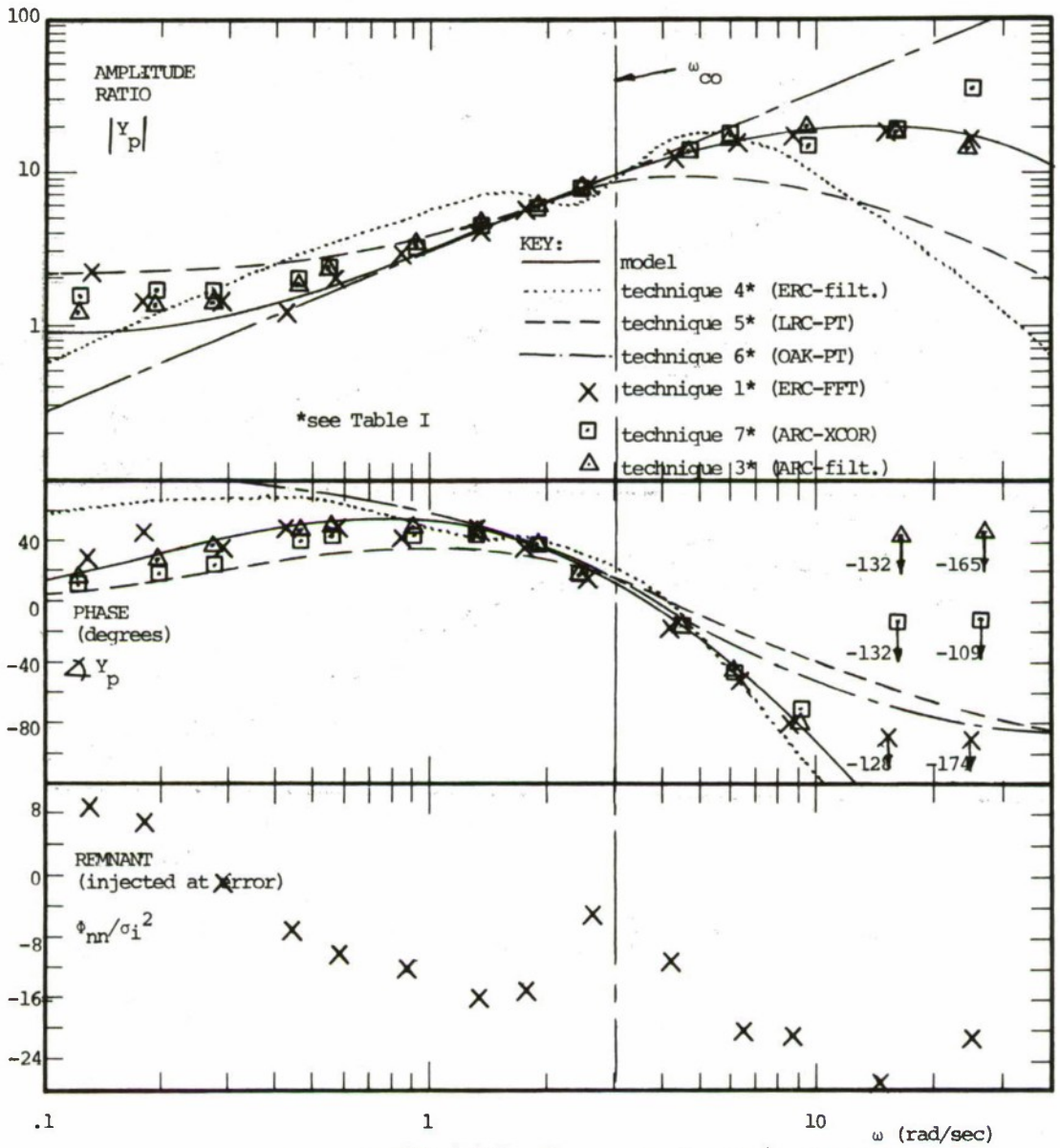
figure 19



RUN NO. 10: $Y_p(s) = .75\,\frac{8-s}{8+s}\,\frac{4.4s+1}{.04s+1}\,\frac{1}{.12s+1}$; $Y_c(s) = 1/s^2$; CLOSED LOOP

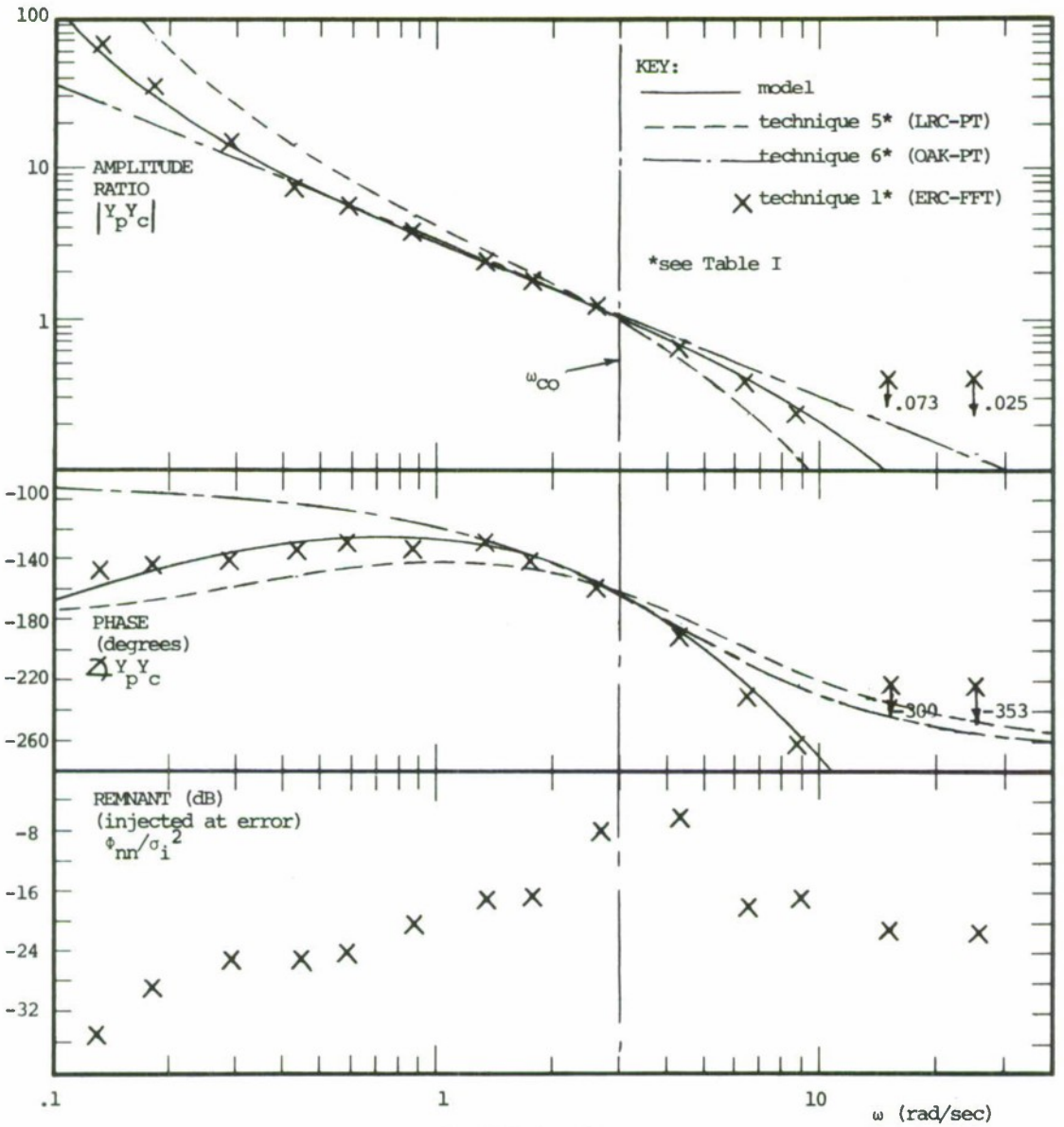
NO NOISE

figure 20



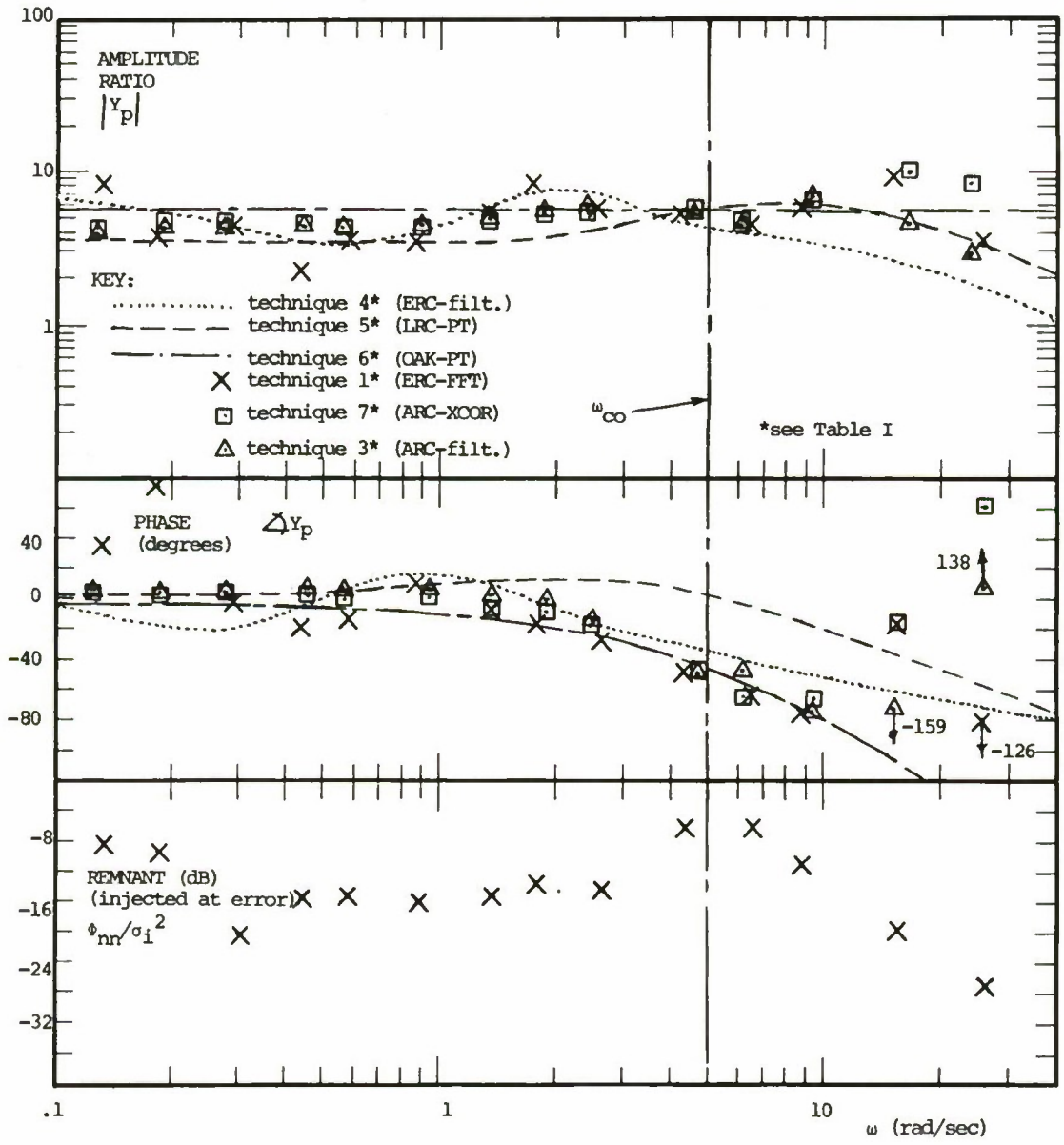
RUN NO. 11: $Y_p(s) = .75 \frac{8-s}{8+s} \frac{4.4s+1}{.04s+1} \frac{1}{.12s+1}$; $Y_c(s) = 1/s^2$; CLOSED LOOP

figure 21



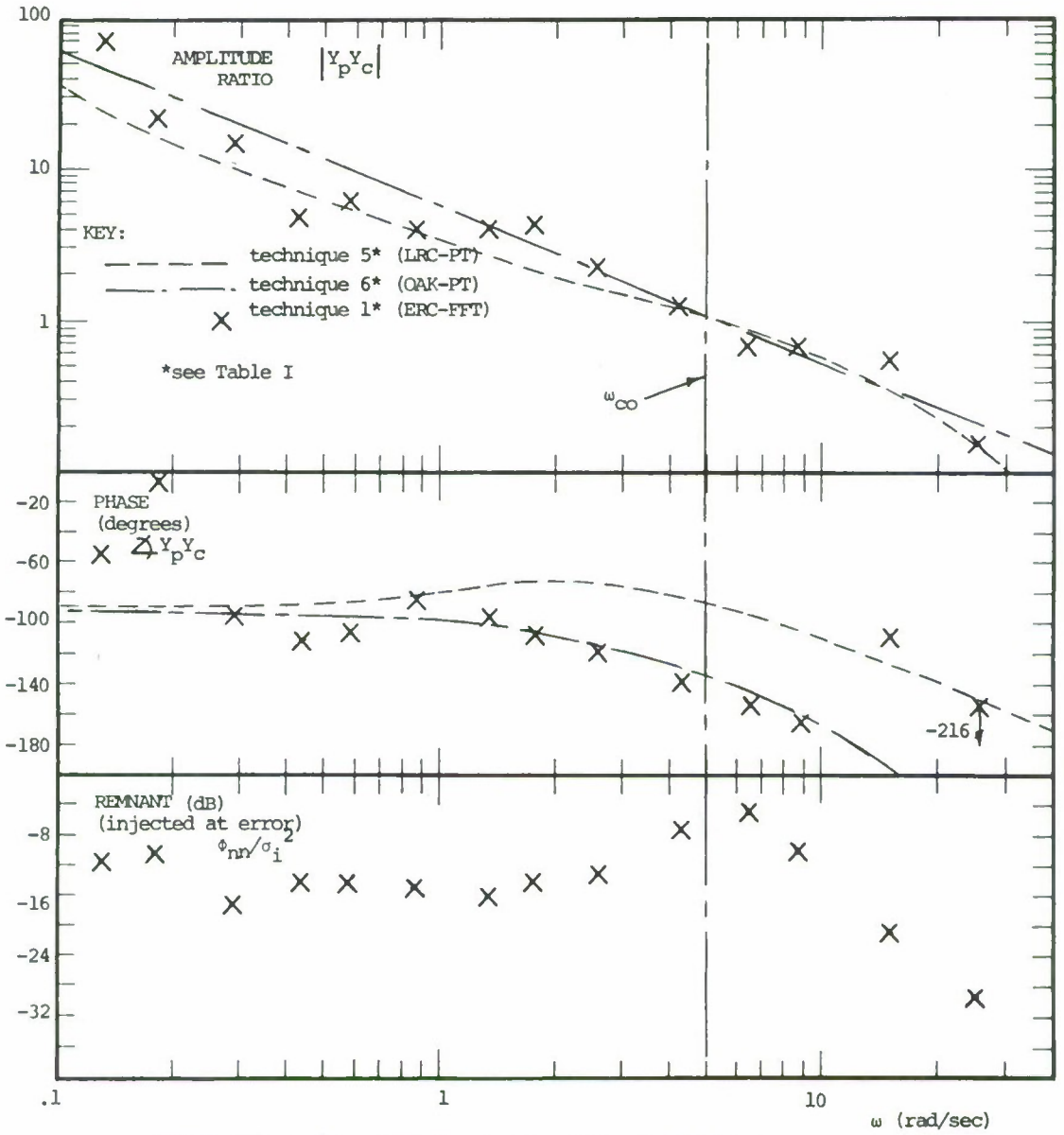
RUN NO. 11: $Y_p(s) = .75\,\frac{8-s}{8+s}\,\frac{4.4s+1}{.04s+1}\,\frac{1}{.12s+1}$; $Y_c(s) = 1/s^2$; CLOSED LOOP

FIGURE 22



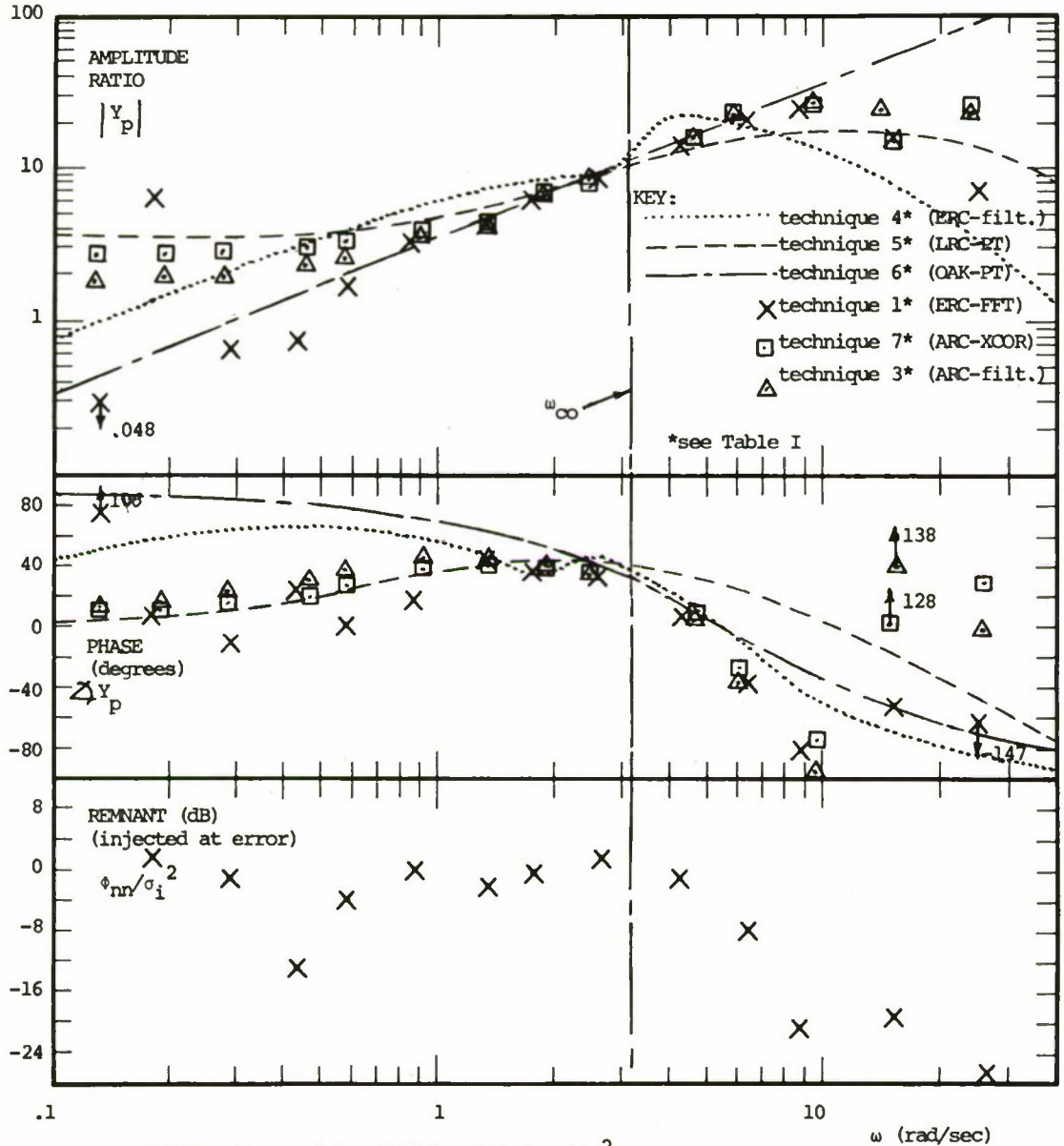
RUN NO. 12: $Y_p(s)$ = PILOT1; $Y_c(s)$ = 1/s; CLOSED LOOP

figure 23



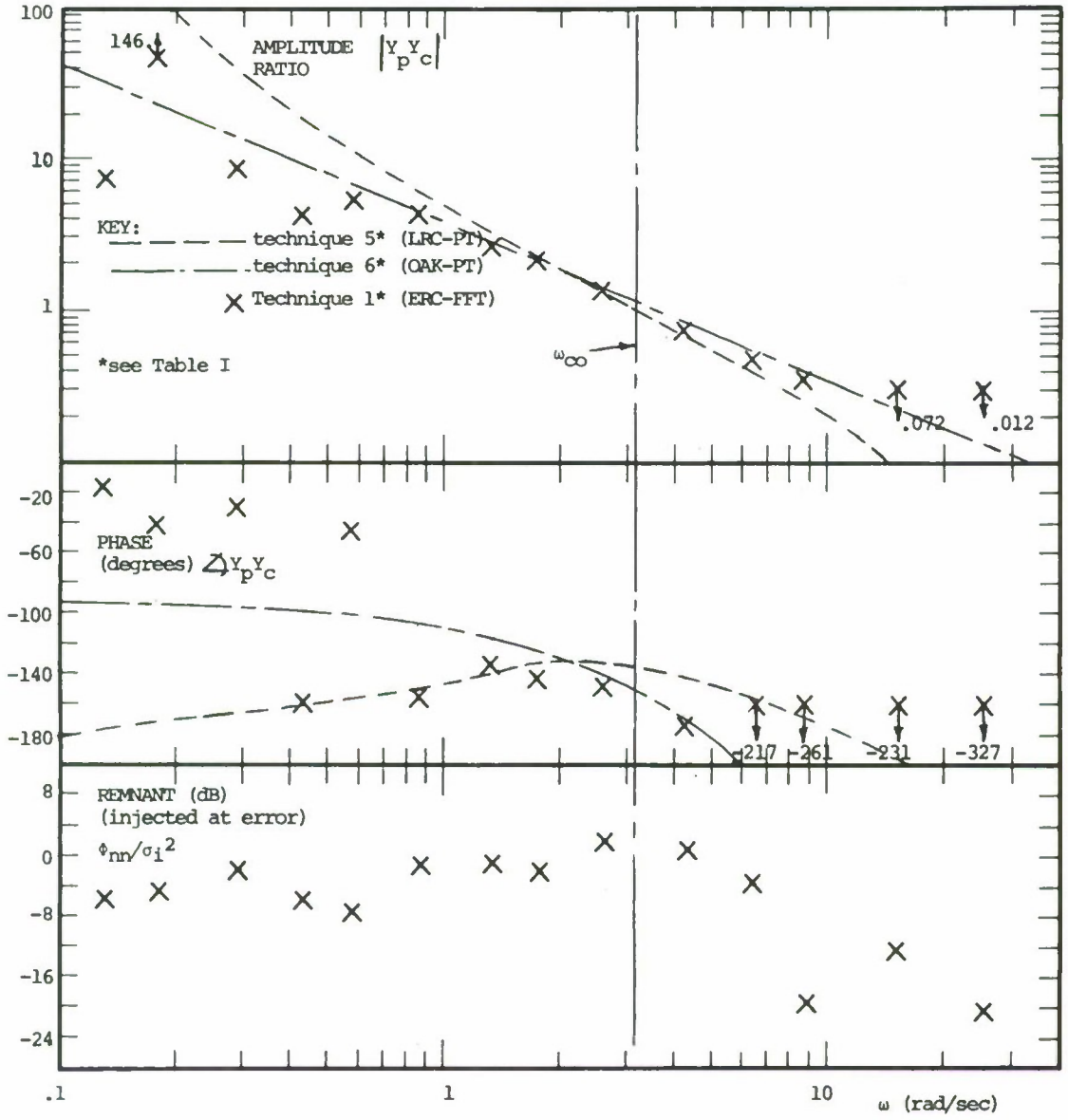
RUN NO. 12: $Y_p(s)$ = PILOT1; $Y_c(s)$ = 1/s; CLOSED LOOP

figure 24



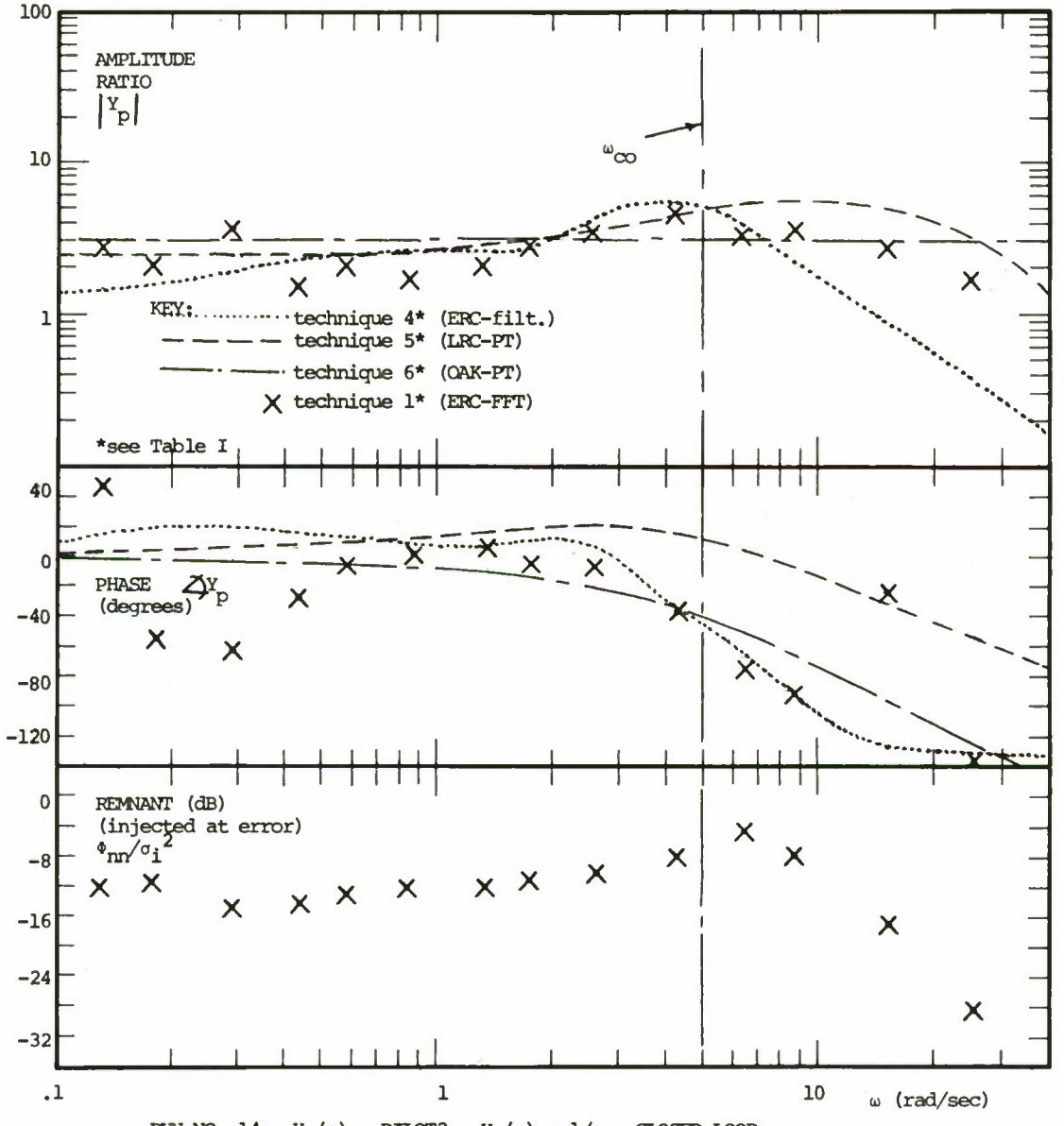
RUN NO. 13: $Y_p(s)$ = PILOT1; $Y_c(s) = 1/s^2$; CLOSED LOOP

figure 25



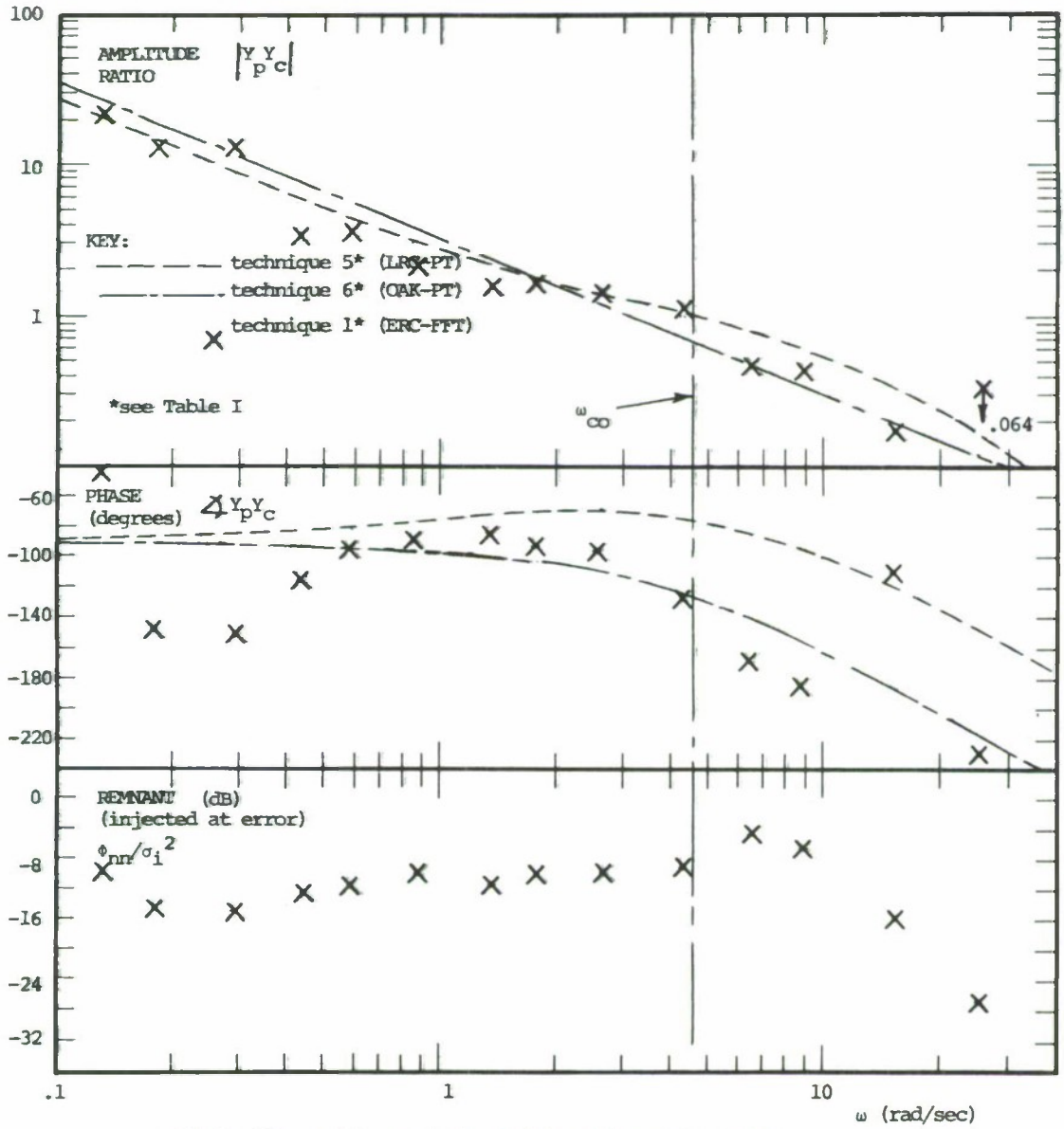
RUN NO. 13: $Y_p(s)$ = PILOT1; $Y_c(s) = 1/s^2$; CLOSED LOOP

figure 26



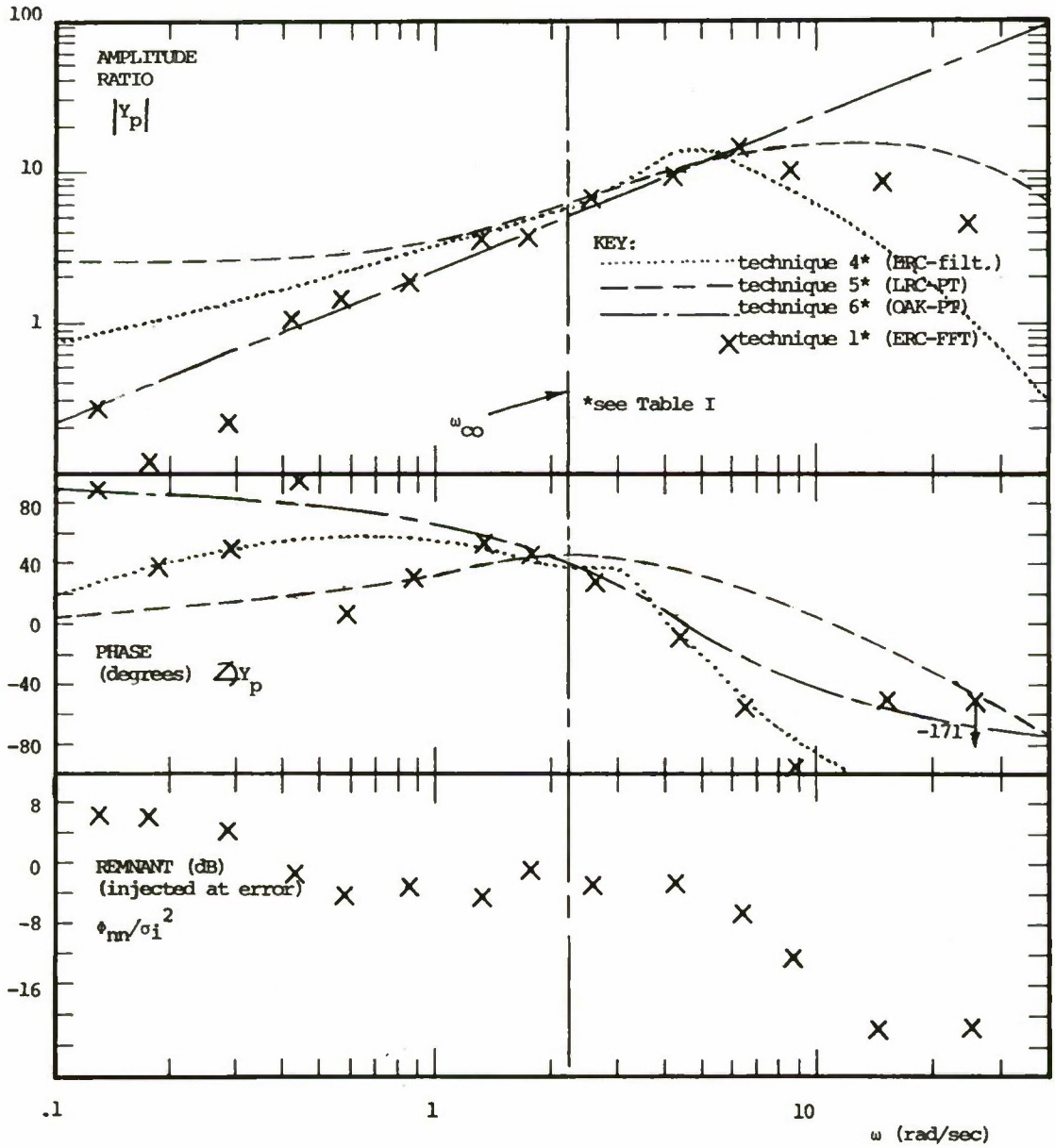
RUN NO. 14: $Y_p(s)$ = PILOT2; $Y_c(s)$ = 1/s; CLOSED LOOP

figure 27



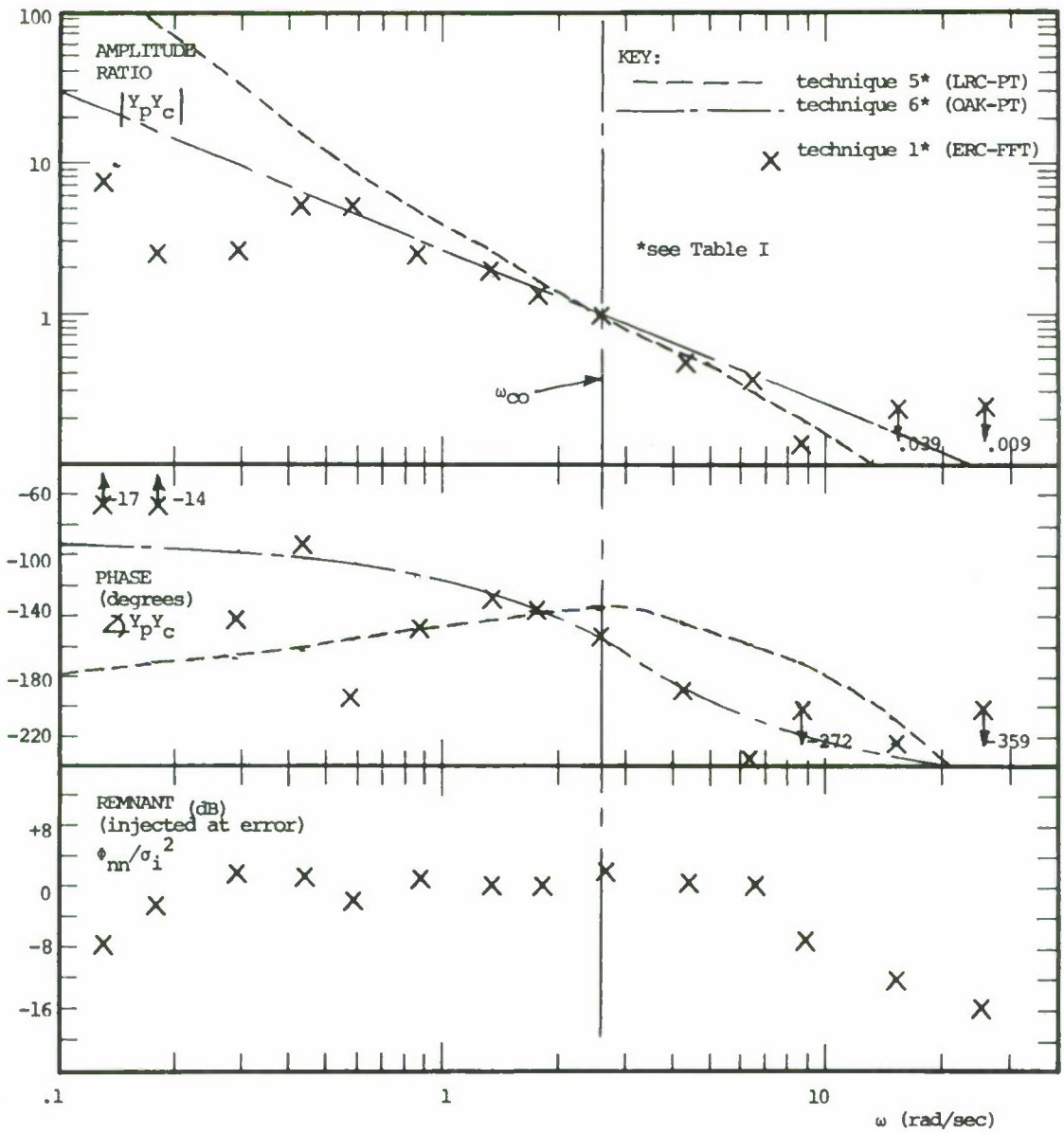
RUN NO. 14: $Y_p(s)$ = PILOT2; $Y_c(s)$ = I/s; CLOSED LOOP

figure 28



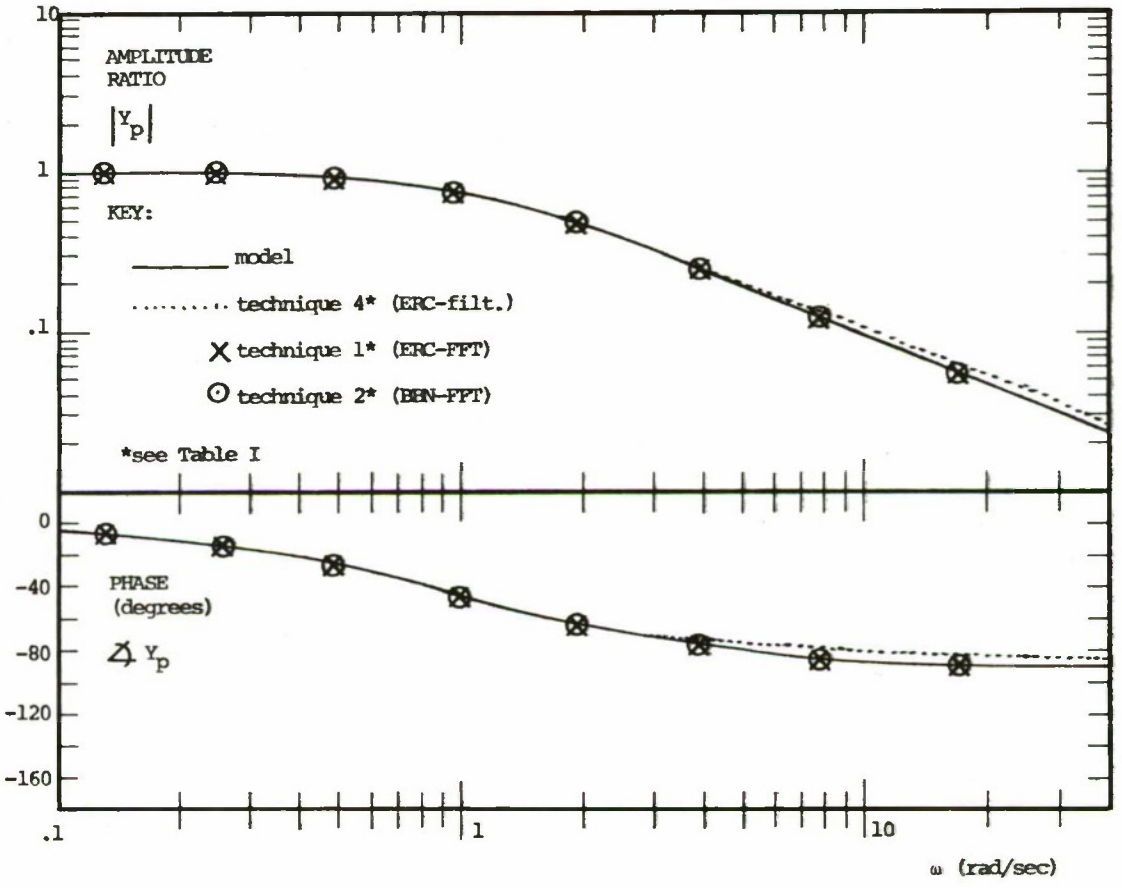
RUN NO. 15: $Y_p(s)$ = PILOT2; $Y_c(s) = 1/s^2$; CLOSED LOOP

figure 29



RUN NO. 15: $Y_p(s)$ = PILOT2; $Y_c(s) = 1/s^2$; CLOSED LOOP

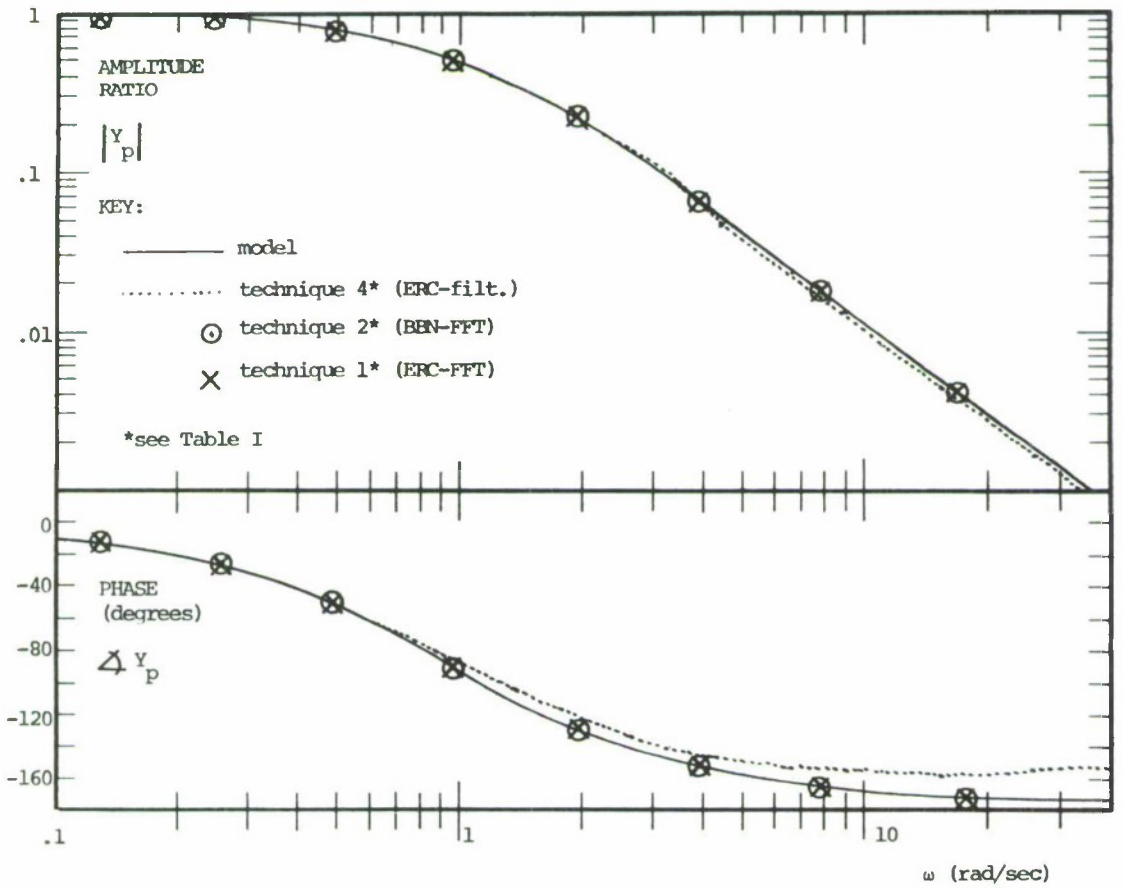
figure 30



RUN NO. 16: $Y_p(s) = (1/s+1)$;

OPEN LOOP NO NOISE

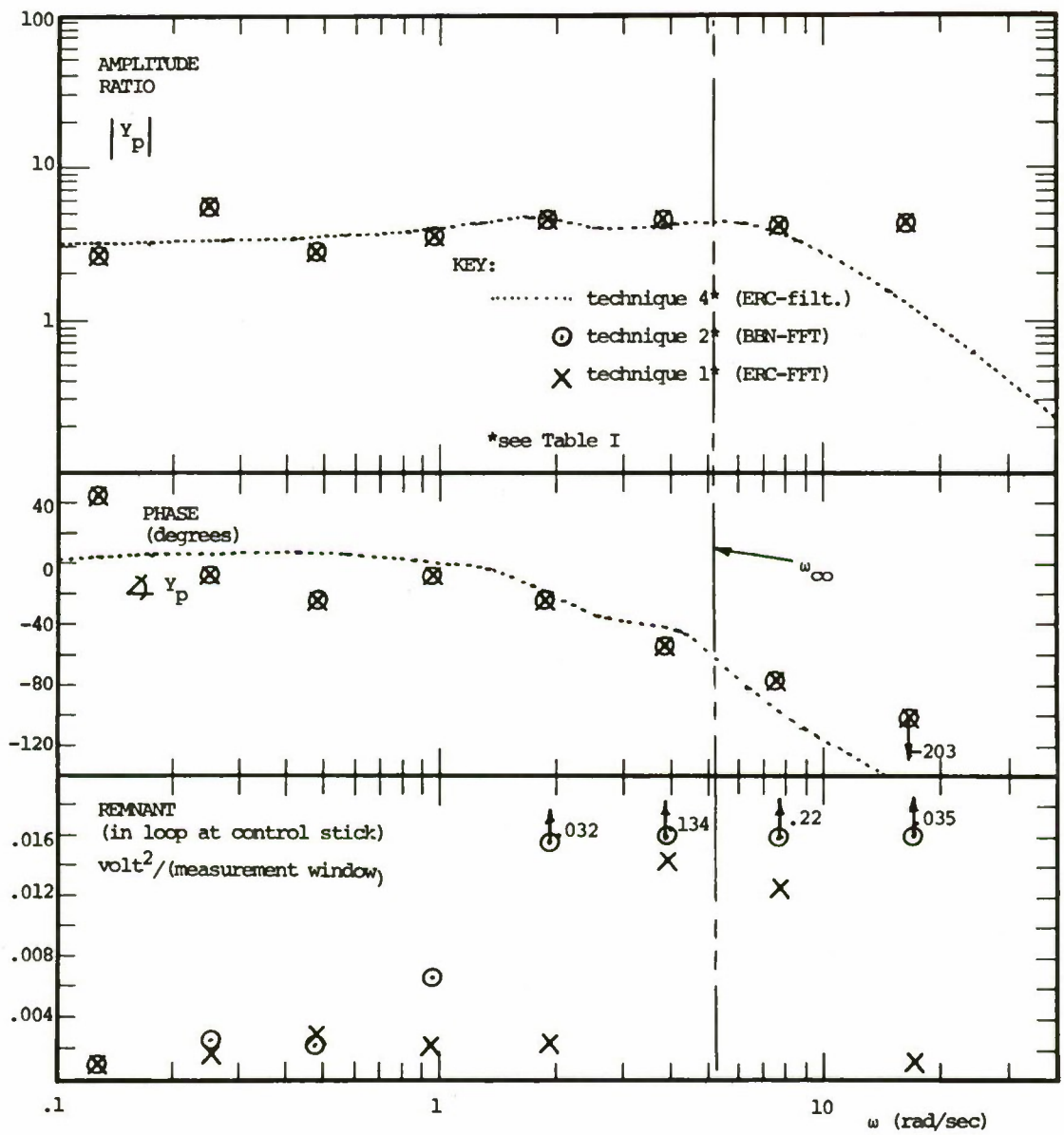
figure 31



RUN NO. 16: $Y_p(s) = (1/s+1)^2$;

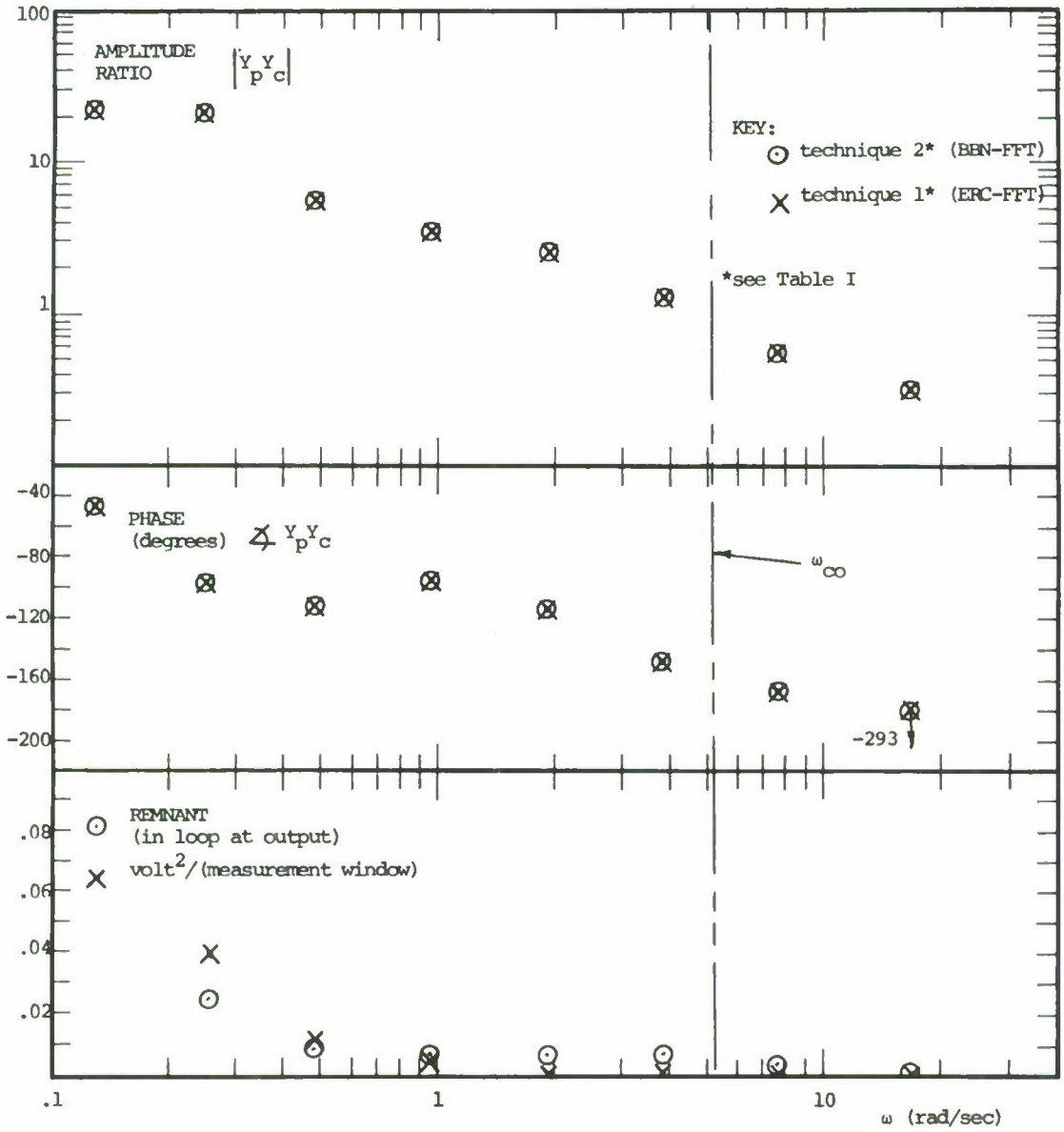
OPEN LOOP NO NOISE

figure 32



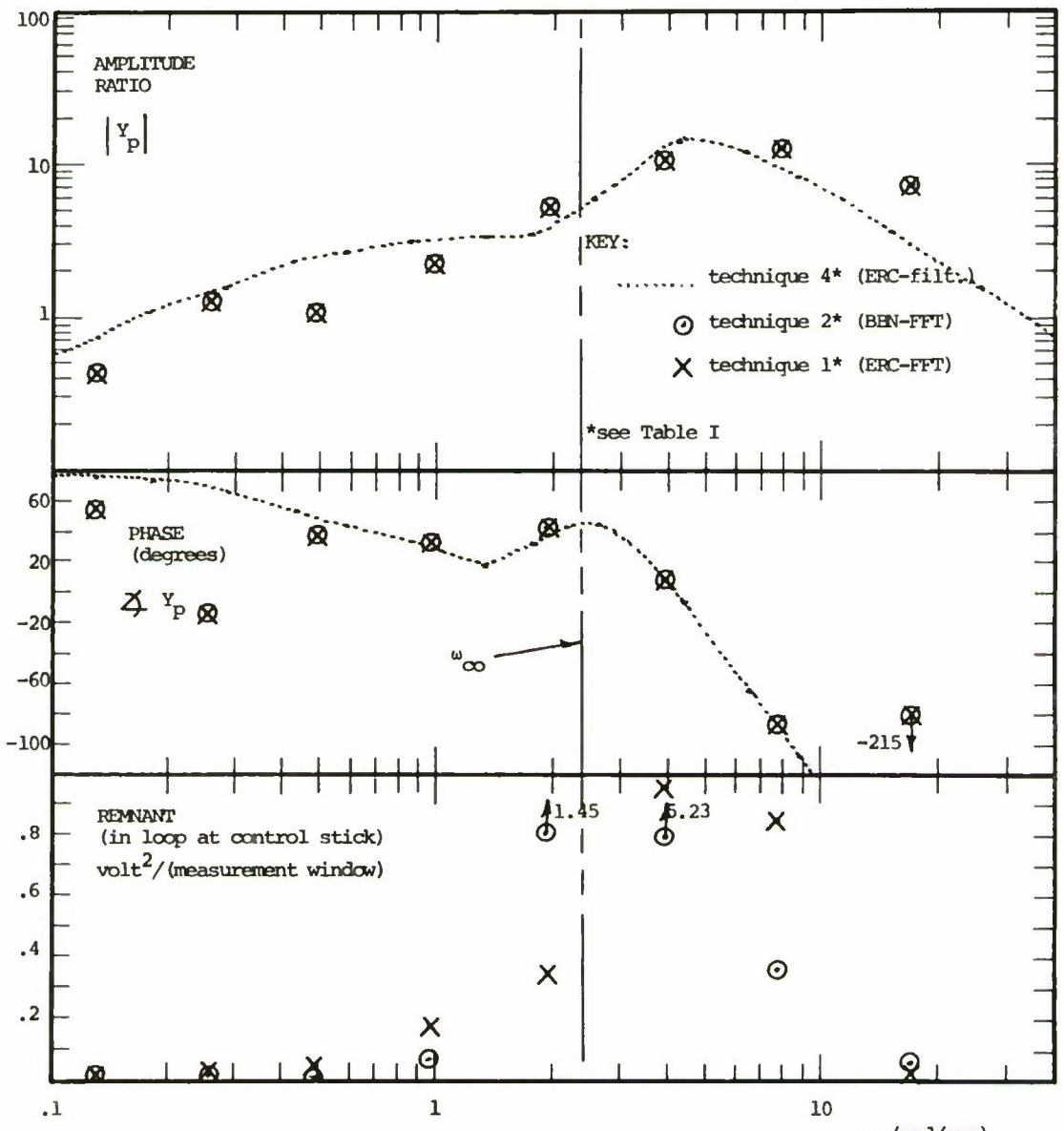
RUN NO. 17: $Y_p(s)$ = PILOT; $Y_c(s)$ = 1/s; CLOSED LOOP

figure 33



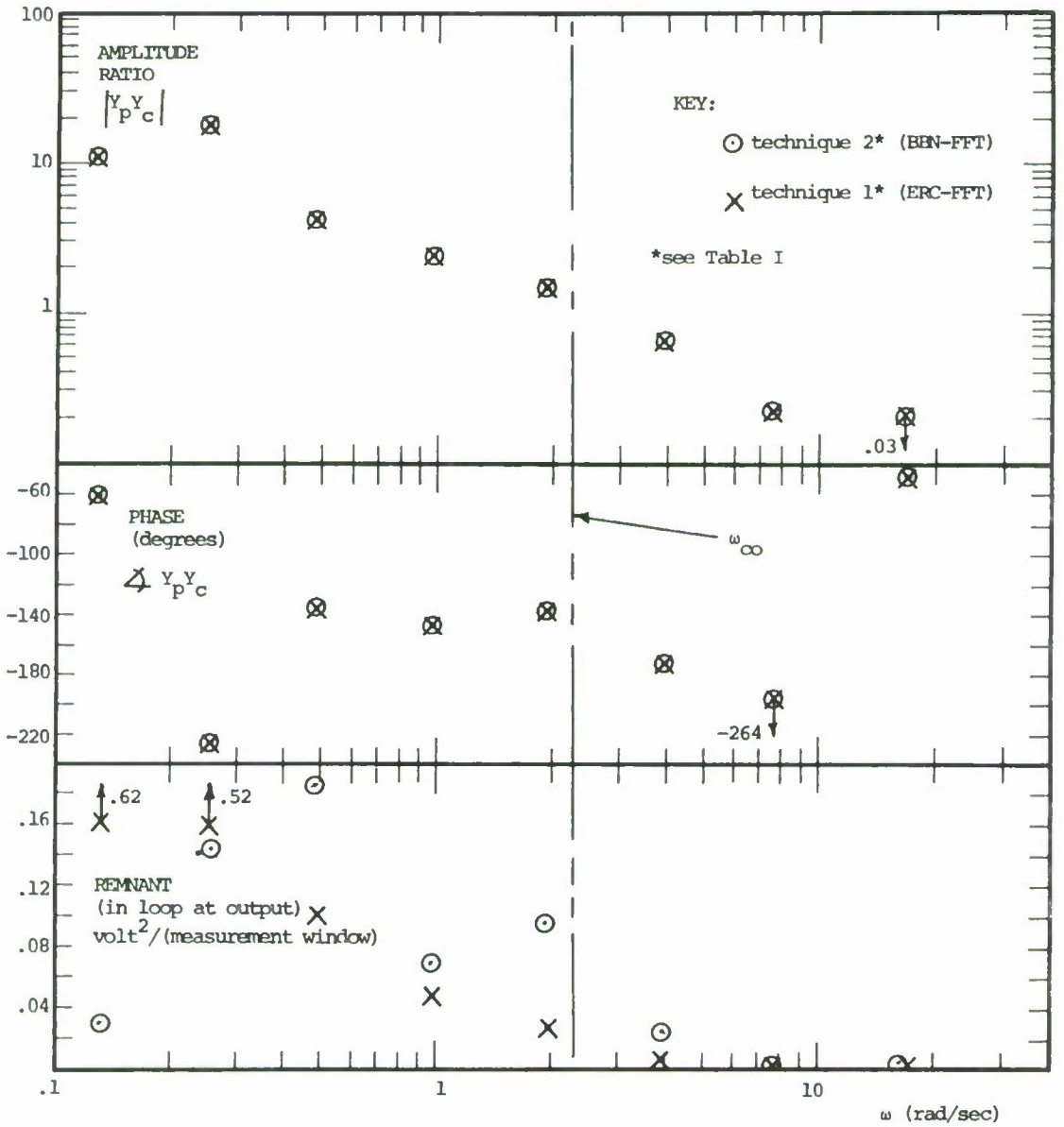
RUN NO. 17: $Y_p(s)$ = PILOT; $Y_c(s)$ = 1/s; CLOSED LOOP

figure 34



RUN NO. 18: $Y_p(s)$ = PILOT; $Y_c(s) = 1/s^2$; CLOSED LOOP

figure 35



RUN NO. 18: $Y_p(s)$ = PILOT; $Y_c(s) = 1/s^2$; CLOSED LOOP

figure 36

A LOOK AT PILOT MODELING TECHNIQUES AT LOW FREQUENCIES

Lawrence W. Taylor, Jr.
NASA Langley Research Center
Hampton, Virginia

ABSTRACT

Pilot models obtained using frequency domain techniques for compensatory tracking sometime exhibit considerable phase lag and excessive scatter at low frequencies, while corresponding models using a time domain analysis do not. It has been suggested, however, that it is because of the limited memory time allowed in the time domain analysis that neither the phase lag nor the scatter in the estimates are evident. It is the purpose of this paper to provide evidence which will help resolve this question.

An analysis of the errors in the estimates of the pilot model characteristics at low frequencies shows the time domain method to be superior but to require a much longer memory time than is usually used. A reanalysis of actual data using a sufficiently long memory time, however, failed to indicate the phase lag indicated by the frequency domain analysis.

INTRODUCTION

Pilot models obtained using frequency domain techniques for single-axis, compensatory tracking sometime exhibit considerable phase lag attendant with large variability in both amplitude ratio and phase, at low frequencies. The sources of the variability or scatter in the estimates of the pilot model characteristics are twofold: (1) the constraint of the limited length of data restricts the number of complete periods for the lowest frequencies and thereby reduces the averaging effect, and (2) the ability of the pilot to control particularly well at low frequencies causes the signal amplitude to become very small and thereby make the effect of noise more pronounced. The phase lag at low frequencies indicated by frequency domain results has caused McRuer et al. in reference 1, and Newell and Smith in reference 2, to empirically fit the results with a time delay made a function of frequency. In this way the amplitude ratio would not be affected. McRuer et al. recognize this device to be only an approximation to something similar to a lag-lead which they use in their "precision crossover model."

It is disturbing, however, to see the large amount of variability, particularly in the results which indicate the pilot to have large phase lag at low frequencies. For example, in reference 2, comparisons are given of the results of three independent frequency domain analyses of the same data which show scatter in the amplitude ratio of as much as an order of magnitude and in phase angle as much as 300°, at the lowest frequency. Similar examples of other

investigators made it clear that there is a problem of determining accurate and reliable estimates of pilot model characteristics at low frequencies.

In reference 3, Taylor shows that time domain results are more consistent and indicate considerably less phase lag at low frequencies than frequency domain results. It has been suggested, however, that it is because of the limited memory time allowed in the time domain analysis that neither the phase lag nor the variable of the estimates are evident. The evidence necessary to resolve this question has been lacking.

The purpose of this paper is to explore these questions, that is, to determine the effects of truncation in both frequency and time domain methods, and to indicate the sensitivity to signals and noise at low frequencies. These analyses will be based on both mathematical models and actual data obtained as part of the joint NASA-USAF-Cornell Aeronautical Laboratory human response studies (ref. 2) involving the T-33 variable-stability airplane and ground simulators.

SYMBOLS

c	pilot output (control deflection), inches
$\overline{c^2}$	mean square or total power of c, inches^2
E	error matrix
e	error, radians
$F[\ ]$	Fourier transform
h_p	impulse response of pilot, inches/degree
I	identity matrix
i	input (external disturbance function), degrees
i_1	lowest frequency input amplitude. $\omega = \omega_1 = 0.157$ radians/sec
K	maximum value of k
k	index for frequency difference
$\mathcal{L}^{-1}$	inverse Laplace transform
M	maximum value of m
m	index for the argument of h_p
N	maximum value of n

n	index for time
$R_{xy}(\tau)$	cross-correlation function
r	remnant signal of pilot model (control deflection), inches
s	Laplace variable
T	total record length, seconds
T_R	limit value of the argument of $R_{xy}(\tau)$, seconds
t	time, seconds
W_k	weighting function
$Y_c(j\omega)$	controlled element transfer function, degrees/inch
$Y_p(j\omega)$	pilot describing function, inches/degree
ρ_a	average linear coherence
σ_r	standard deviation of the remnant, inches
$\sigma_{F[r]}$	standard deviation of the Fourier transform of the remnant
τ	argument of h_p, seconds
ϕ	phase angle, rad
$\Phi_{xy}(j\omega)$	cross-spectral density of $x(t)$ and $y(t)$
$\Phi_{xx}(\omega)$	power-spectral density of $x(t)$
τ_{max}	maximum memory time of the pilot model, seconds
ω	frequency, radians/second
ω_r	cutoff frequency of remnant
ω_1	lowest input frequency, $\omega_1 = 0.157$ rad/sec
$\|\|$	amplitude ratio
$\measuredangle$	phase angle, degrees or radians
$\hat{}$	estimate

Matrix notation:

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

The subscript 1 denotes relationship to the lowest input frequency, ω_1.

DESCRIPTION OF EXPERIMENT

In the classical single-axis, compensatory tracking experiment, the pilot is asked to minimize the error, displayed to him by an oscilloscope, television screen, or meter by manipulating a controller. The controller deflection is sent to an analog computer which computes the response of the controlled element and adds to it the input disturbance function 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 which are later processed to obtain the model of the pilot. Similar experiments have been performed in flight. The data analyzed in this paper were collected as part of the joint NASA-USAF-Cornell Aeronautical Laboratory human response studies (ref. 2) involving the T-33 variable-stability airplane and ground simulators.

DISCUSSION

Example Results

Although the phenomena of phase lag and large variability of pilot models have been observed for a variety of controlled element characteristics, we will limit our discussion to a single-axis task where $Y_c = \frac{20}{s}$ and for which there are considerable data. Figure 1 shows a comparison of results (ref. 2) of Newell and Smith for three independent analyses using frequency domain techniques for the same data. It is evident from the figure that the scatter, and therefore the uncertainty of the estimates of the pilot model amplitude ratio and phase angle, is largest at the lowest frequency. It is also evident that the average phase angle is on the order of 90° although the large amounts of scatter make it impossible to be precise.

Similar results (ref. 1) of McRuer et al. presented in figure 2 show the phase lag and the scatter, particularly in phase angle at the lowest frequencies.

Shown on the same figure is a comparison of these results with the "extended crossover model" also of reference 1. The "extended crossover model" differs from its earlier form by an additional factor which provides phase lag at low frequencies without affecting the amplitude ratio. The reader is warned in reference 1 that this model is only an approximation which should not be applied at frequencies beyond those shown (i.e., less than $\omega = 0.157$ rad/sec). If one did not heed this warning and tried to mechanize such a model, difficulties would be encountered because of instability caused by phase lag exceeding 180° with an open loop gain greater than unity at low frequencies. In the same reference McRuer et al. suggest a more reasonable "precision crossover model," in which the phase lag at low frequency is attributed to a lag lead.

In contrast to these frequency domain results, time domain results are shown in reference 3 to have relatively little variability and negligible phase lag at low frequency. The same comparison is shown in figure 3. It has been suggested, however, that the apparent advantage of the time domain technique might be caused by a lack of sensitivity resulting from a memory time which is too short. Because the evidence necessary to resolve this question has been lacking, an analysis is made of both frequency domain techniques and the time domain technique, with regard to errors in the resulting model characteristics at low frequency.

Frequency Domain Techniques

For the purpose of this discussion, we will consider only input disturbance or forcing functions which have discrete spectra and data lengths which are an integral multiple of the periods of the input frequencies analyzed so that the estimate $\hat{Y}_p$ simplifies, as is shown in reference 4, to:

$$\hat{Y}_p(j\omega) = \frac{\Phi_{ic}(j\omega)}{\Phi_{ie}(j\omega)} = \frac{\sum_{k=-k}^{k} W_k F_k^*[i]\, F_k[c]}{\sum_{k=-k}^{k} W_k F_k^*[i]\, F_k[e]} = \frac{F[c]}{F[e]}$$

Provided $F_k[i] = 0$ for $k \neq 0$ and W_k becomes zero before the adjacent input frequency is reached.

Effect of Remnant.- The effect of the remnant on estimates of Y_p will be studied through the use of a statistical model. From the block diagram of figure 4 we can write:

$$F[c] = \frac{Y_p F[i] + F[r]}{1 + Y_p Y_c}$$

$$F[e] = \frac{F[i] - Y_c F[r]}{1 + Y_p Y_c}$$

The frequency domain estimate becomes

$$\hat{Y}_p = \frac{Y_p + \frac{F[r]}{F[i]}}{1 - Y_c \frac{F[r]}{F[i]}}$$

The familiar observation can again be made that as the ratio of the Fourier transform of the remnant to that of the input goes to zero, the estimate $\hat{Y}_p$ goes to its correct value Y_p. But as the remnant to input ratio goes to infinity the estimate $\hat{Y}_p$ becomes $-\frac{1}{Y_c}$. If we consider the ratio $\frac{F[r]}{F[i]}$ as a random vector whose phase has a uniform distribution over $\pm\pi$ radians and whose amplitude has a gaussian distribution, we can determine the probability density of $\hat{Y}_p$. First, let us establish an approximate relationship between the power of the remnant and the variance of its Fourier transform.

Consider the remnant to be an unbiased random variable with a gaussian distribution and to have a spectrum that is constant at frequencies below a cutoff frequency, ω_r of 5 rad/sec (based on ref. 5). Let the standard deviation of the remnant be denoted by σ_r. We can write:

$$\sigma_r^2 = \frac{1}{\pi}\int_0^{\omega_r} \Phi_{rr}(\omega)\, d\omega = \frac{\omega_r \Phi_{rr}}{\pi} \quad \text{for} \quad 0 < \omega < \omega_r$$

The expected value of $F^*[r]\, F[r]$ or $\sigma^2_{F[r]}$ is related to Φ_{rr} by:

$\frac{\sigma^2_{F[r]}}{T} = \Phi_{rr}$ which is based on the relationship between the Fourier transform of a signal and its power spectral density given in reference 4, for example. Solving for $\sigma_{F[r]}$ in the above expressions we get:

$$\sigma_{F[r]} = \sqrt{\frac{2\pi T}{\omega_r}}\, \sigma_r$$

If $i(t) = i_1 \sin(\omega_1 t + \phi_1)$ the amplitude of its Fourier transform is:

$$\left|F[i_1]\right|_{\omega = \omega_1} = \left|\int_0^T i_1 \sin(\omega_1 t + \phi_1)(\cos \omega_1 t - j \sin \omega_1 t)\, dt\right| = \frac{i_1 T}{2}$$

provided the half period, $\frac{\pi}{\omega_1}$, is contained an integral number of times in the record, T.

We can now express

$$\frac{\sigma_{F[r]}}{|F[i_1]|} = 2\sqrt{\frac{\pi}{\omega_r T}}\,\frac{\sigma_r}{i_1} = 0.145\,\frac{\sigma_r}{i_1}$$

for $T = 120$ sec and $\omega_r = 5$ rad/sec.

It is now possible to compute the characteristics of $\hat{Y}_p$ as a function of the remnant to input ratio, $\frac{\sigma_r}{i_1}$. A Monte Carlo technique was used to perform these calculations for which 100 samples of $\frac{F[r]}{F[i]}$ were selected at random from a population having the probability density function discussed earlier. The 15, 50, and 85 percentile values of the resulting values of $\hat{Y}_p$ were then noted. The procedure was repeated for several values of $\frac{\sigma_r}{i_1}$ until the function presented in figures 5(a) and (b) generated. If the distribution of $\hat{Y}_p$ were gaussian, the interval indicated would correspond to $\pm 1\sigma$.

The scatter is seen to increase for increasing values of the remnant to input ratio, σ_r/i_1, until a point is reached where the estimate is strongly biased toward $-\frac{1}{Y_c}$ after which the scatter diminishes, until $\hat{Y}_p$ becomes $-\frac{1}{Y_c}$ exactly.

In order to establish a typical value of the remnant to input ratio, σ_r/i_1, it is useful to use the approximate relationship for $\overline{c^2}$:

$$\overline{c^2} \approx \frac{1}{2}\sum_{k=1}^{10} i_k^2 \left|\frac{1}{Y_c(i\omega_k)}\right|^2 \approx 0.00067\, i_1^2$$

Because it is assumed that the component of $\overline{c^2}$ due to the remnant is negligible, it is good for relatively small levels of remnant power only.

Since ρ_a^2 is defined (ref. 1) as:

$$\rho_a^2 = 1 - \frac{\sigma_r^2}{\overline{c^2}} \approx 1 - \frac{\sigma_r^2}{0.00067\, i_1^2}$$

solving for $\frac{\sigma_r}{i_1}$ we get

$$\frac{\sigma_r}{i_1} \approx 0.015 \quad \text{for} \quad \rho_a^2 = 0.65 \quad \text{(ref. 3)}$$

For this value of remnant to input ratio figures 5(a) and (b) would indicate that the scatter (70 percent of the values) of the amplitude ratio would range over a factor of 2 and phase angle would vary $\pm 35^{\circ}$. There would be negligible bias in either amplitude ratio or phase angle. If, however, only 30 seconds of data were analyzed, instead of 2 minutes, the function would be shifted to the left by a factor of 2 and the scatter would more than double, and the average amplitude ratio would be biased to a value about 20 percent larger than the true value of 0.07. The reason for this effect of record length can be seen by examining the previously developed expression which showed that the variance of the Fourier transform of the remnant was proportional to the record length times the variance of the remnant.

The effect of Y_c on the variance can been seen by examining the expression for $\hat{Y}_p$. For the example considered $|Y_c| >> \left|\frac{1}{Y_p}\right|$ so that the effect of $F[r]/F[i]$ on $\hat{Y}_p$ is caused almost entirely by the term in the denominator for typical values of the remnant to input ratio. The variance of $\hat{Y}_p$, then, will increase proportionally to Y_c^2 so for $Y_c = \frac{K}{s}$ the standard deviation will be inversely proportional to ω. For the example considered, the scatter will be about 1.6 times as much at ω_1 as for ω_2.

Effect of Truncation.- Another source of error for frequency domain estimates of Y_p which first determine cross-correlation functions, is the truncation involved in computing the Fourier transform of the cross-correlation functions. Let us consider the following forms of Y_p and Y_c.

$$Y_p = K_p \frac{\left(\frac{1}{T_k} + s\right)}{\left(\frac{1}{T_k'} + s\right)} = 0.07 \frac{(0.3 + s)}{(0.05 + s)} \qquad Y_c = \frac{50}{s}$$

The form of Y_p selected is based on the "precision crossover model" of reference 1 and whose frequency response is shown in figure 6. We will again be concerned with effects of $\hat{Y}_p$ at low frequencies but the effect of the remnant will now be ignored. Let us look at the estimated amplitude ratio and phase angle of the following estimate as the value of T_r is limited.

$$\hat{Y}_p = \frac{\hat{\Phi}_{ic}(j\omega)}{\hat{\Phi}_{ie}(j\omega)} = \frac{\int_{-T_R}^{T_R} R_{ic}(\tau)^{-j\omega\tau} d\tau}{\int_{-T_R}^{T_R} R_{ie}(\tau)^{-j\omega\tau} d\tau}$$

Let us further consider only the lowest input frequency so that

$$i = i_1 \sin \omega_1 t$$

$$e(t + \tau) = \frac{i_1}{\left|1 + Y_p Y_c\right|} \sin \left(\omega_1 t + \measuredangle \frac{1}{1 + Y_p Y_c} + \omega_1 \tau \right)$$

$$c(t + \tau) = i_1 \left| \frac{Y_p}{1 + Y_p Y_c} \right| \sin \left(\omega_1 t + \measuredangle \frac{Y_p}{1 + Y_p Y_c} + \omega_1 \tau \right)$$

$$R_{ie}(\tau) = \frac{1}{T} \int_0^T i(t)\, e(t+\tau)\, dt = \frac{i_1^2}{\left|1 + Y_p Y_c\right|} \left[\frac{\cos \left(\measuredangle \frac{1}{1 + Y_p Y_c} + \omega_1 \tau \right)}{2} \right.$$

$$\left. + \frac{1}{4\omega_1 T} \left(\cos \left(\measuredangle \frac{1}{1 + Y_p Y_c} + \omega_1 \tau \right) \sin 2\omega_1 T + \sin \left(\measuredangle \frac{1}{1 + Y_p Y_c} \omega_1 \tau \right) (1 - \cos 2\omega_1 T) \right) \right]$$

If T is chosen to contain an integral number of periods of ω_1 (or half periods)

$$R_{ie}(\tau) = \frac{i_1^2}{2\left|1 + Y_p Y_c\right|} \cos \left(\measuredangle \frac{1}{1 + Y_p Y_c} + \omega_1 \tau \right)$$

Similarly for $R_{ic}(\tau)$

$$R_{ic}(\tau) = \frac{i_1^2}{2\left|1 + Y_p Y_c\right|} \cos \left(\measuredangle \frac{Y_p}{1 + Y_p Y_c} + \omega_1 \tau \right)$$

The cross-spectral density (ref. 4) for both would be:

$$\Phi_{ie}(j\omega_1) = \frac{i_1^2}{2}\left|\frac{Y_p}{1 + Y_p\,Y_c}\right|\left[T_R\, e^{j\measuredangle \frac{Y_p}{1 + Y_p\,Y_c}} + \frac{\sin 2\omega_1 T_R}{2\omega_1}\, e^{-j\left(\measuredangle \frac{Y_p}{1 + Y_p\,Y_c}\right)}\right]$$

$$\Phi_{ic}(j\omega_1) = \frac{i_1^2}{2\left|1 + Y_p\,Y_c\right|}\left[T_R\, e^{j\measuredangle \frac{1}{1 + Y_p\,Y_c}} + \frac{\sin 2\omega_1 T_R}{2\omega_1}\, e^{-j\left(\measuredangle \frac{1}{1 + Y_p\,Y_c}\right)}\right]$$

which gives:

$$\hat{Y}_p(j\omega_1) = Y_p(j\omega_1)\left[\frac{1 + e^{-j2\measuredangle \frac{Y_p}{1 + Y_p\,Y_c}} \frac{\sin 2\omega_1 T_R}{2\omega_1 T_R}}{1 + e^{-j2\measuredangle \frac{1}{1 + Y_p\,Y_c}} \frac{\sin 2\omega_1 T_R}{2\omega_1 T_R}}\right]$$

Figure 7 shows how the estimate of $\hat{Y}_p$ changes as T_R is varied. It is important to note that if T_R contains an integral number of half periods of the frequency analyzed, there is no bias on $\hat{Y}_p$ caused by the truncation.

Time Domain

Now that we have examined sources of error in frequency domain estimates of Y_p, let us turn our attention to time domain estimates. By the term "time domain pilot model" (refs. 3 and 4), we mean the impulse response function $\hat{h}_p(\tau)$ which minimizes

$$\overline{r^2} = \left(c - \int_0^{T_M} \hat{h}_p(\tau)\, e(t - \tau)\, d\tau\right)^2$$

and its Fourier transform, $\hat{Y}_p$.

Using matrix notation we can express the least-squares solution as

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

where:

$$\underline{c} = \begin{pmatrix} c_M \\ c_{M+1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ c_N \end{pmatrix} \qquad \underline{\hat{h}}_p = \begin{pmatrix} \hat{h}_p(\tau_1) \\ \hat{h}_p(\tau_2) \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \hat{h}_p(\tau_M) \end{pmatrix} \qquad E = \begin{bmatrix} e_M & \cdots & e_2 & e_1 \\ e_{M+1} & & & \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ e_N & & & e_{N-M+1} \end{bmatrix}$$

The estimate $\hat{Y}_p$ is then obtained by taking the Fourier transform of $\hat{h}_p$.

$$\hat{Y}_p = F\left[\hat{Y}_p\right]$$

Again a Monte Carlo approach was taken in which data were generated using the following model and repeatedly analyzed to obtain values of $\hat{Y}_p$ for several values of the remnant to input ratio.

$$Y_p = 0.07$$

$$Y_c = \frac{50}{s}$$

$$i = \sum_{k=1}^{10} i_k \sin \omega_k t$$

(Frequencies and amplitudes of the input are the same as those in refs. 1, 3.)

$$r = \text{white noise filtered by } \frac{1}{1 + 0.2s}$$

Figure 8 shows the results of the analysis in terms of confidence limits (15 and 85 percentiles) on the amplitude ratio and phase angle of $\hat{Y}_p$ for the lowest input frequency as a function of the remnant to input ratio, σ_r/i_1. The results shown in figure 8 are significantly different from the corresponding results for the frequency domain estimates for Y_p in figure 5.

The variability indicated by the vertical bars (15 to 85 percentile) is smaller for the time domain estimates of Y_p except for extreme values of σ_r/i_1 where the frequency domain estimate becomes $\frac{1}{Y_c}$. The bias is less in both amplitude ratio and phase angle for the estimates in the time domain in general but are about the same for $\left|\hat{Y}_p\right|$ at low value of σ_r/i_1.

Effect of Limited Memory.- The effect of limited memory of the time domain pilot model is assumed to be that of truncating the impulse response function. The Y_p used to study this effect is that used previously, the frequency response for which is given in figure 6.

$$Y_p(s) = 0.07\,\frac{(s + 0.3)}{(s + 0.05)}$$

It is instructive to look at the impulse response $h_p(\tau)$ for this model in figure 9.

$$h_p(\tau) = \mathcal{L}^{-1}\left(Y_p(s)\right) = 0.0175\,e^{-0.05\tau} + 0.07\,s(\tau)$$

The describing function $\hat{Y}_p(j\omega)$ equals the Fourier transform of $h_p(\tau)$.

The effect of truncation can be studied by examining the finite integral

$$\hat{Y}_p(j\omega) = \int_0^{\tau_{max}} h_p(\tau)\,e^{-j\omega\tau}\,d\tau$$

$$= 0.07\,\frac{(j\omega + 0.3)}{(j\omega + 0.05)} - \frac{0.0175}{(j\omega + 0.05)}\,e^{-(j\omega + 0.05)\,\tau_{max}}$$

$$= Y_p(j\omega) - \frac{0.0175}{(j\omega + 0.05)}\,e^{-(j\omega + 0.05)\,\tau_{max}}$$

This function was used to determine $\left|\hat{Y}_p\right|\Big/\left|Y_p\right|$ and the error in $\measuredangle\,\hat{Y}_p$ which is plotted in figure 10 as a function of the memory time, τ_{max}. It can be seen by looking at figure 10 that if the actual pilot exhibited the characteristics of the example pilot model, a memory time far exceeding the 2.5 to 5 sec usually used (refs. 3 and 4) would be required.

This result would suggest a reanalysis of actual human response data to see if greatly increased values of T_M would result in a $\hat{Y}_p$ that would exhibit lag-lead characteristics.

Reanalysis of Experimental Data.- In figure 11, a $\hat{Y}_p$ for which $\tau_{max} = 40$ sec is compared with earlier results (ref. 3) for $\tau_{max} = 2.5$ sec. Although small differences exist, there is no indication of a lag-lead of the type previously modeled and indicated by frequency domain estimates even though the value of τ_{max} is large enough to allow such results. This result in connection with the greater accuracy indicated for analysis in the time domain would imply that the lag-lead indicated by frequency domain analysis is a result of high variability and bias at low frequency.

CONCLUDING REMARKS

1. Errors in $\hat{Y}_p$ due to remnant and truncation for analysis in both the time and frequency domains are computed for $Y_c = \frac{K}{s}$.

2. Time domain $\hat{Y}_p$ exhibited less variance and less bias than for frequency domain estimates for the same remnant to input ratios.

3. In frequency domain analysis the record length and the range of τ in the cross-correlation function should be chosen to contain an integral number of half periods of the frequencies analyzed.

4. The values of τ_{max} previously used for analysis in the time domain excluded indications of lag-lead characteristics at low frequency.

5. A rerun of data for sufficiently long τ_{max} resulted in little change indicating that the lag-lead indications are due to bias and variability of frequency domain measures.

REFERENCES

1. McRuer, Duane; Graham, Dunstan; Krendel, Ezra; and Reisener, William, Jr.: Human Pilot Dynamics in Compensatory Systems. Theory, Models, and Experiments With Controlled Element and Forcing Function Variations. Tech. Rep. AFFDL-TR-65-15, Air Force Flight Dynamics Lab., Wright-Patterson Air Force Base, July 1965.

2. Newell, Fred D.; and Smith; Harriet J.: Human Transfer Characteristics in Flight and Ground Simulation for a Roll Tracking Task. NASA TN D-5007, February 1969.

3. Taylor, Lawrence W., Jr.: Nonlinear Time-Domain Models of Human Controllers. Fifth Annual NASA-University Conference on Manual Control, MIT, Cambridge, Mass., 1969.

4. Taylor, Lawrence W., Jr.: A Comparison of Human Response Modeling in the Time and Frequency Domains. USC/NASA Conference on Manual Control, Los Angeles, California, March 1-3, 1967, NASA SP-144.

5. Levision, William H.; and Keinman, David L.: A Model for Human Controller Remnant Fourth Annual NASA-University Conference on Manual Control, Ann Arbor, Michigan, March 21-23, 1968, NASA SP-192.

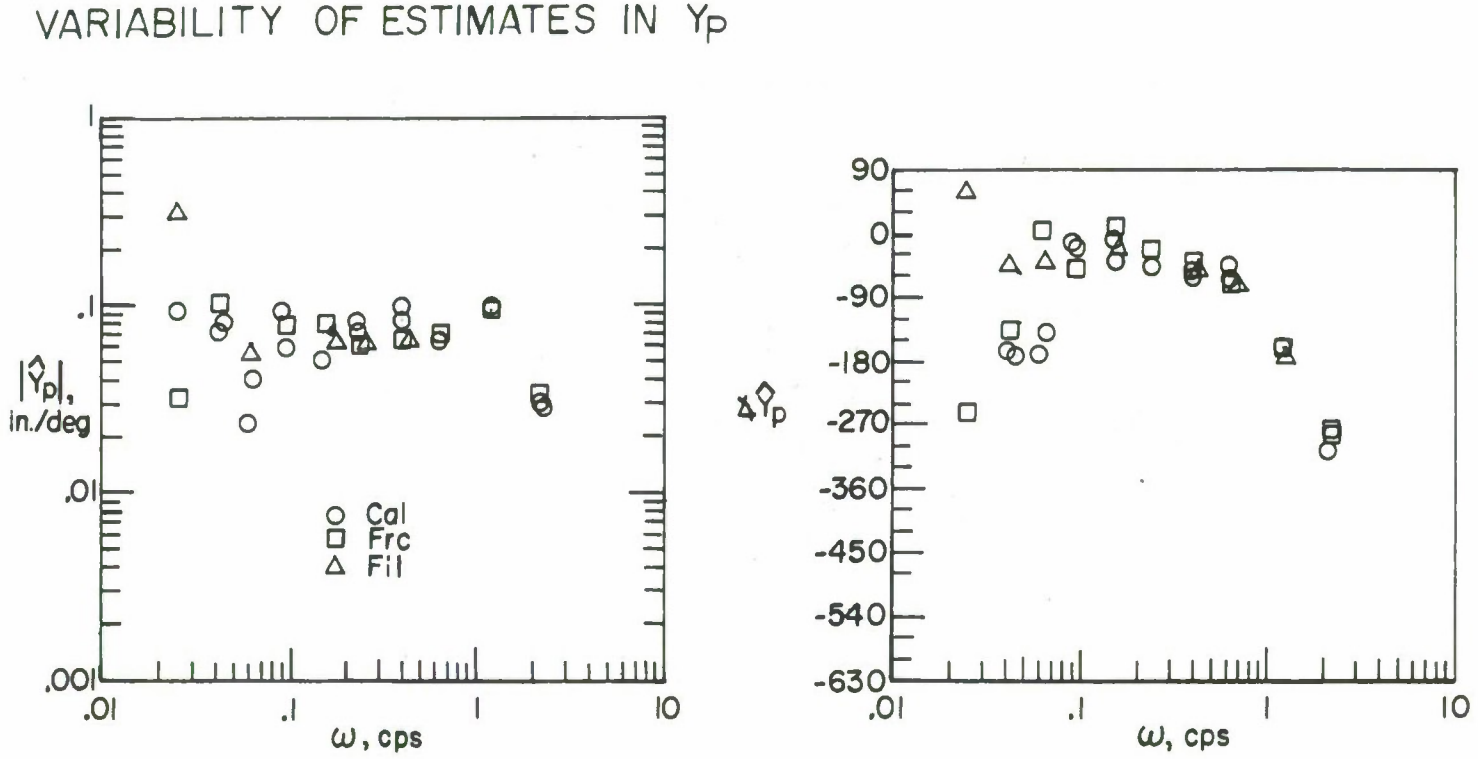


Figure 1.- Comparison of results for three independent analyses in the frequency domain of the same data. Reference 2.

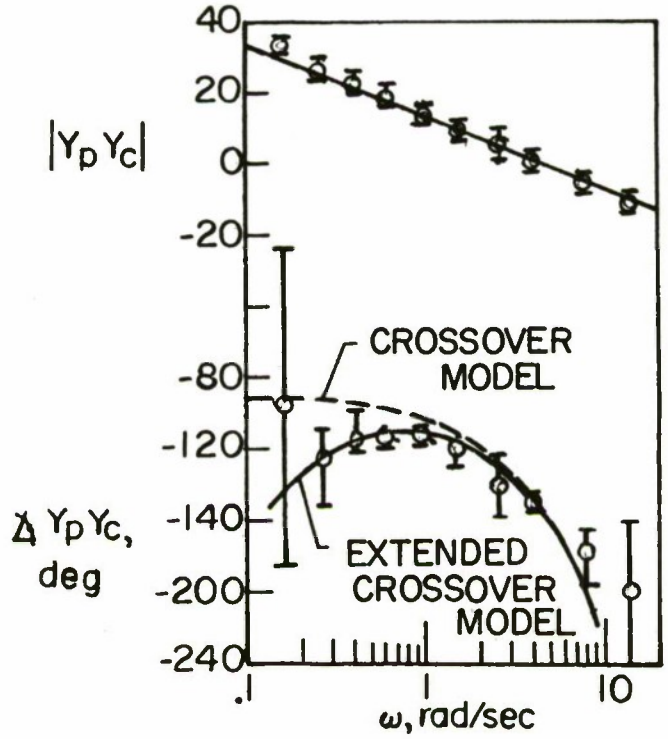


Figure 2.- Results and crossover models of reference 1.

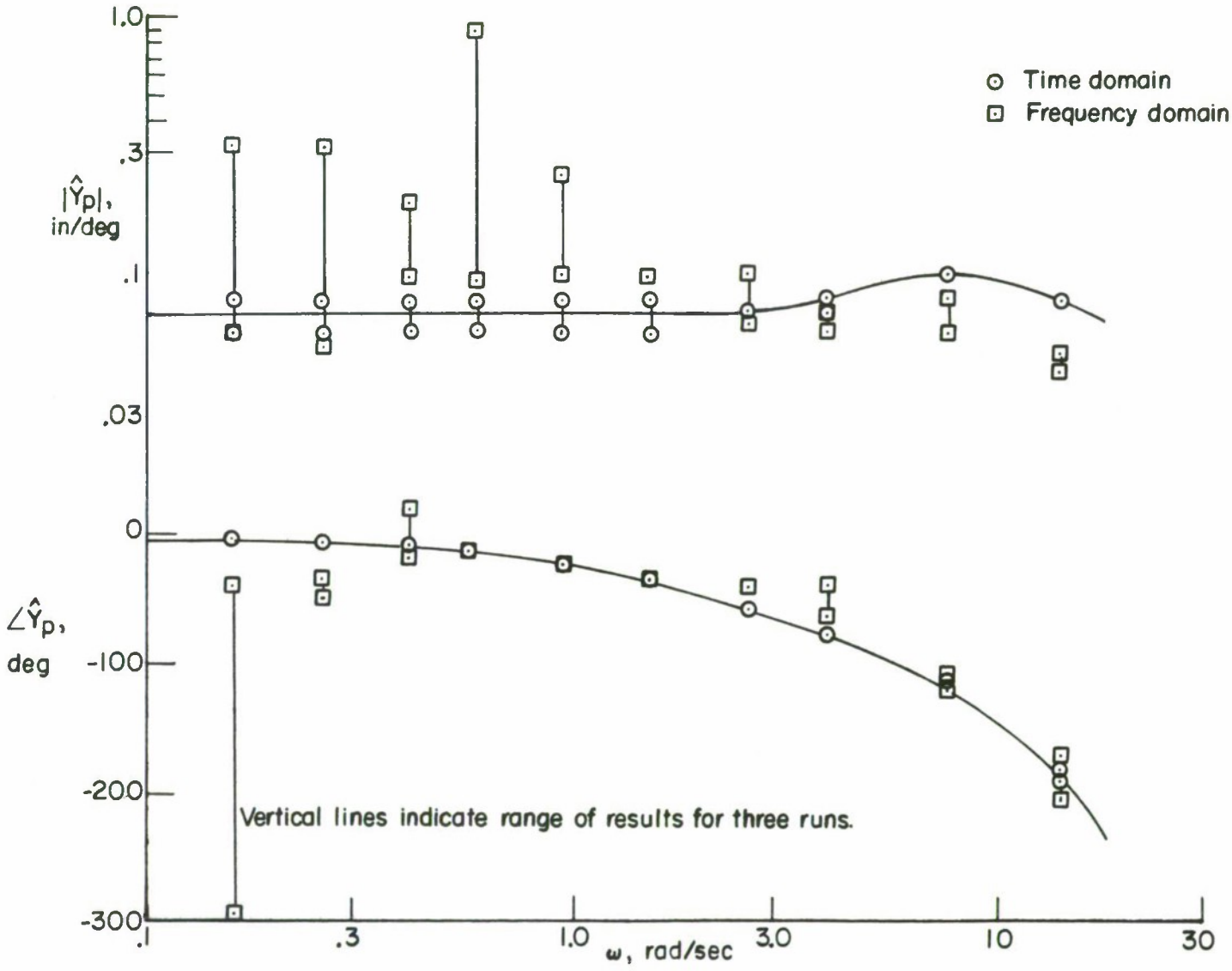


Figure 3.- A comparison of results of analyses in the frequency and time domains. Reference 3.

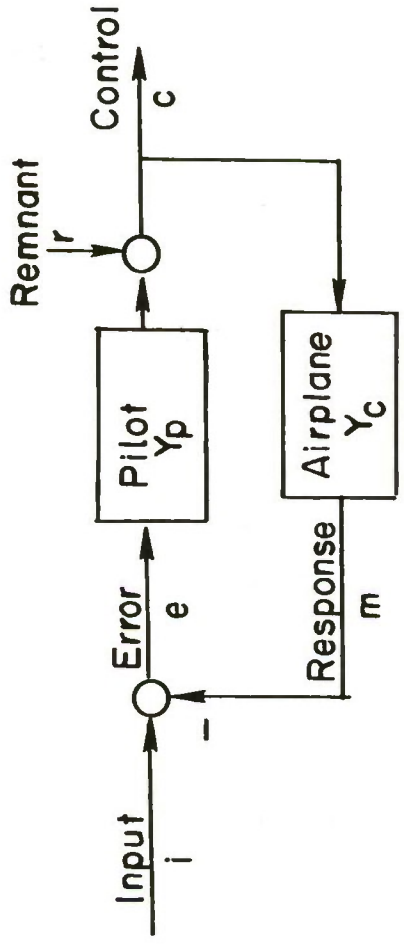
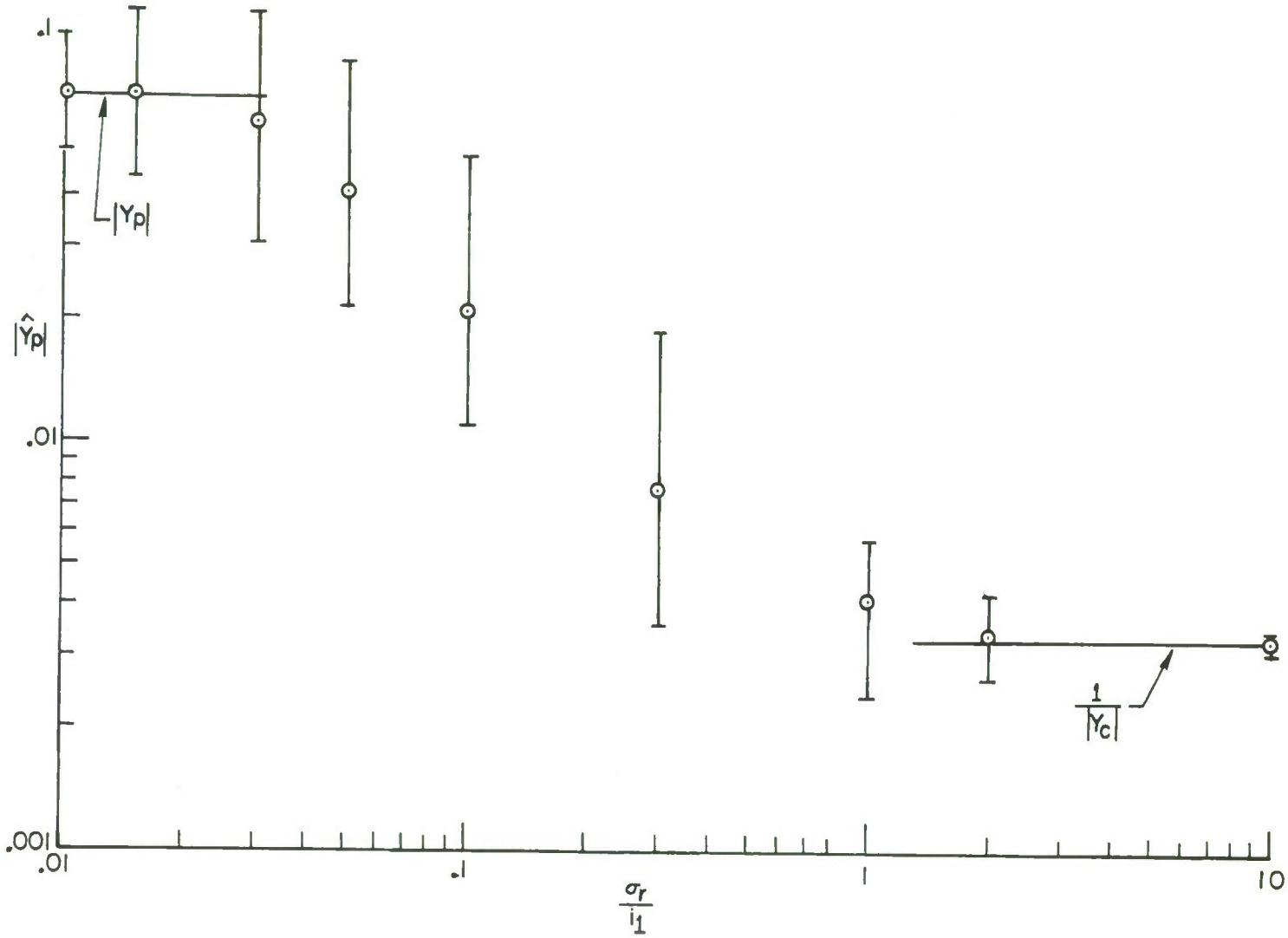
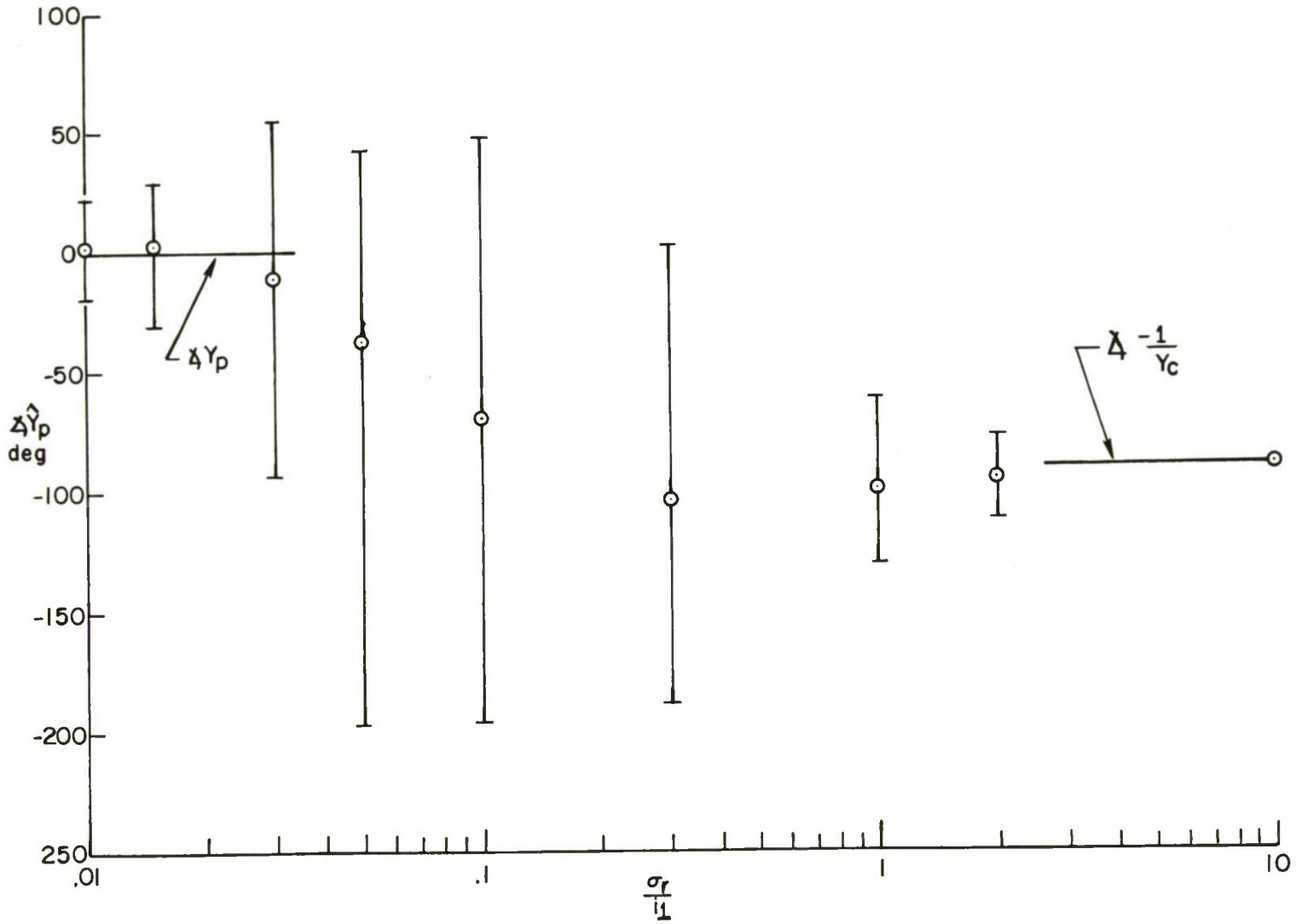


Figure 4.- Block diagram of compensatory tracking task.



(a) Amplitude ratio.

Figure 5.- Effect of remnant to input ratio on frequency domain estimate of Y_p.



(b) Phase angle.

Figure 5.- Concluded.

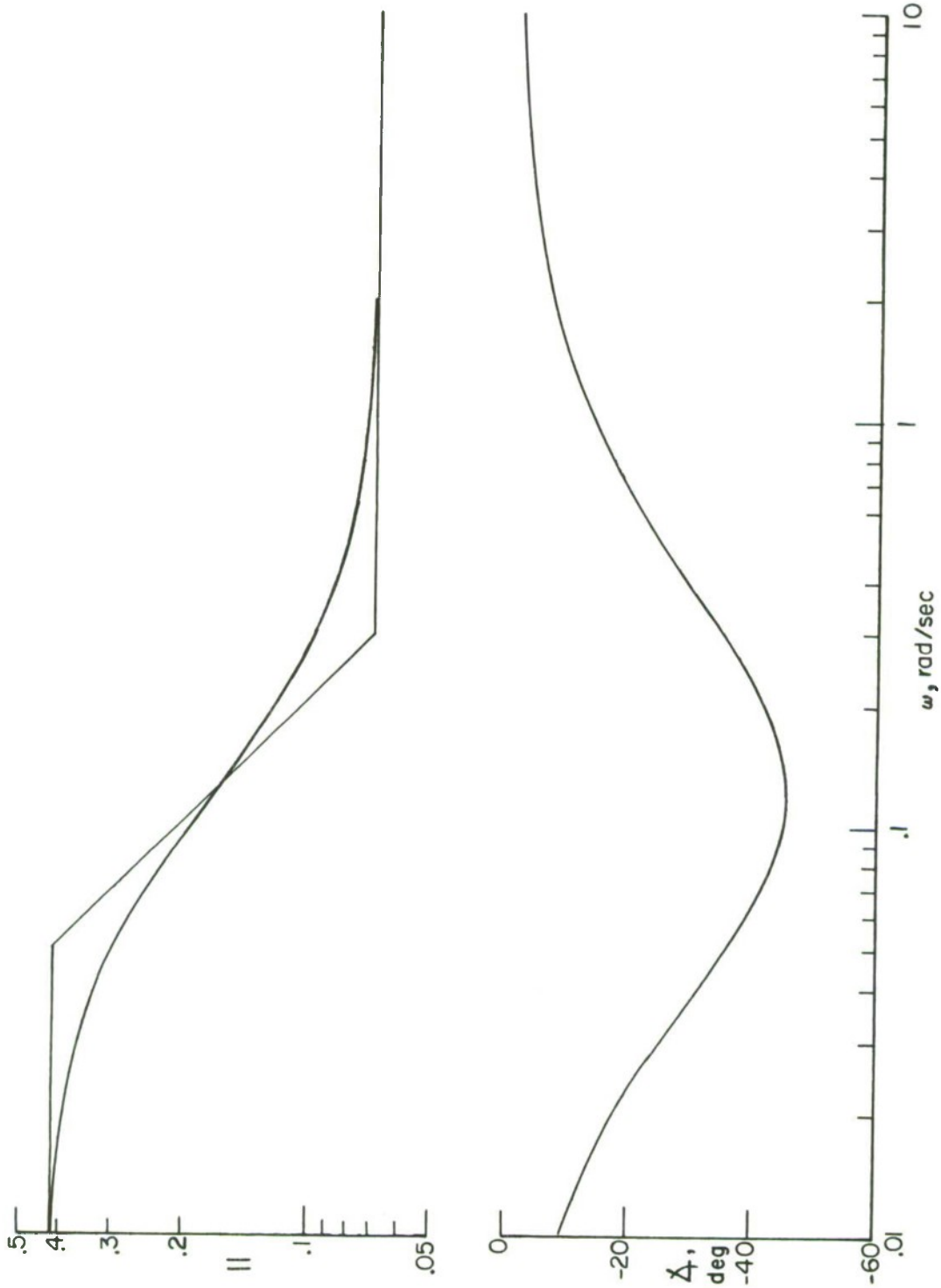


Figure 6.- Frequency response of the pilot model used to study truncation effects.

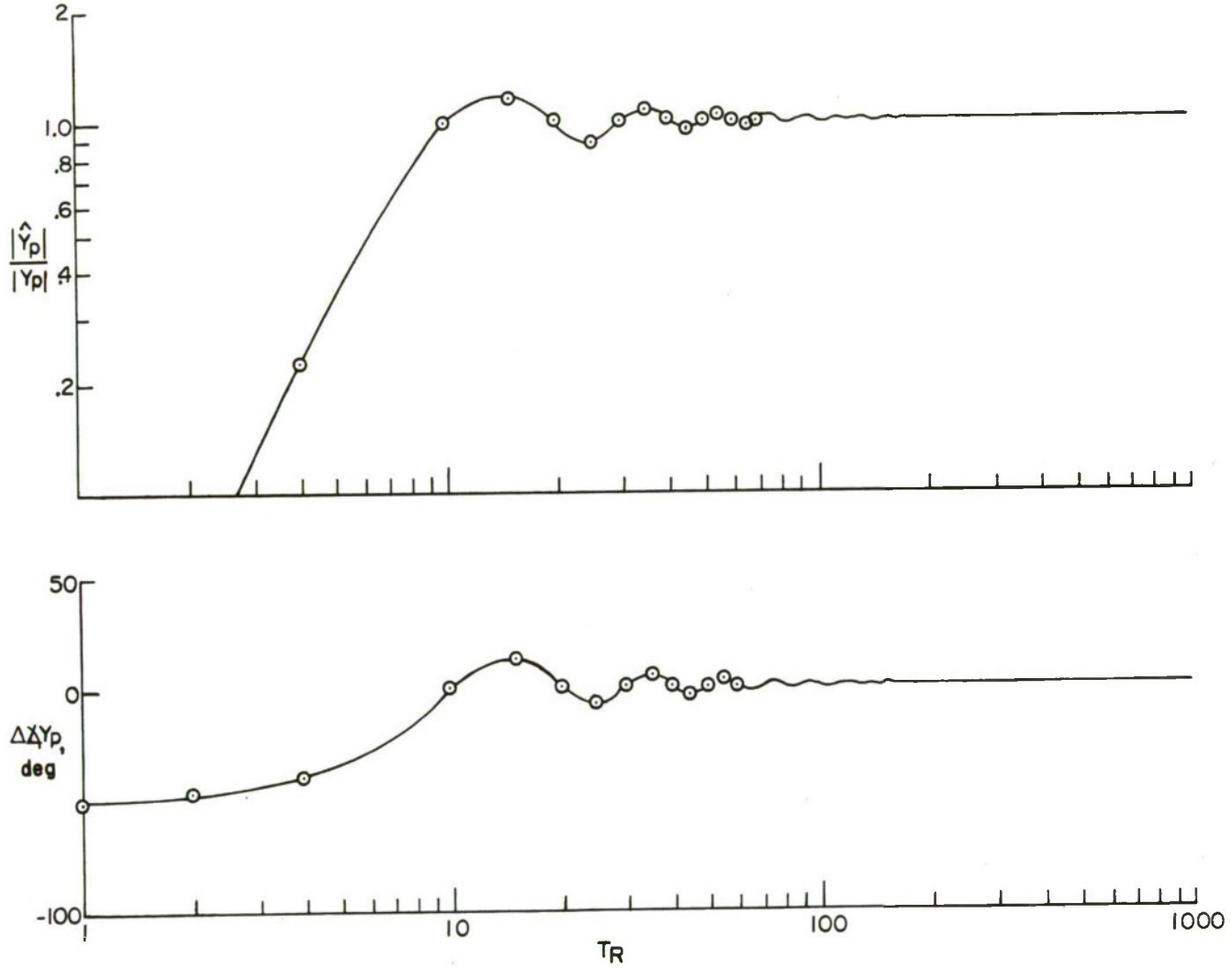


Figure 7.- Effect of truncation of cross-correlation function on Y_p estimate. ($\omega = 0.157$.)

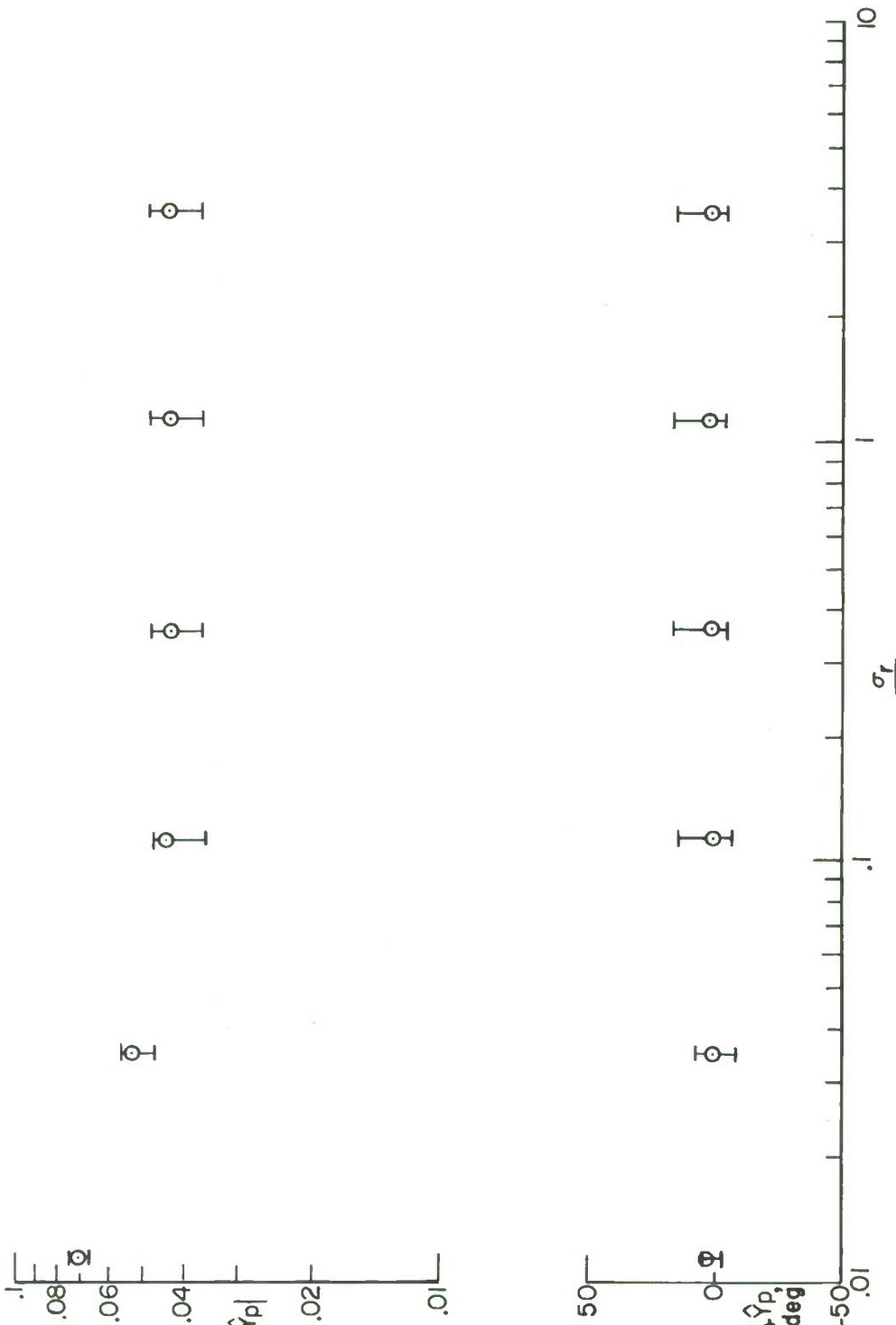


Figure 8.- Effect of remnant to input ratio on time domain estimate of Y_p.

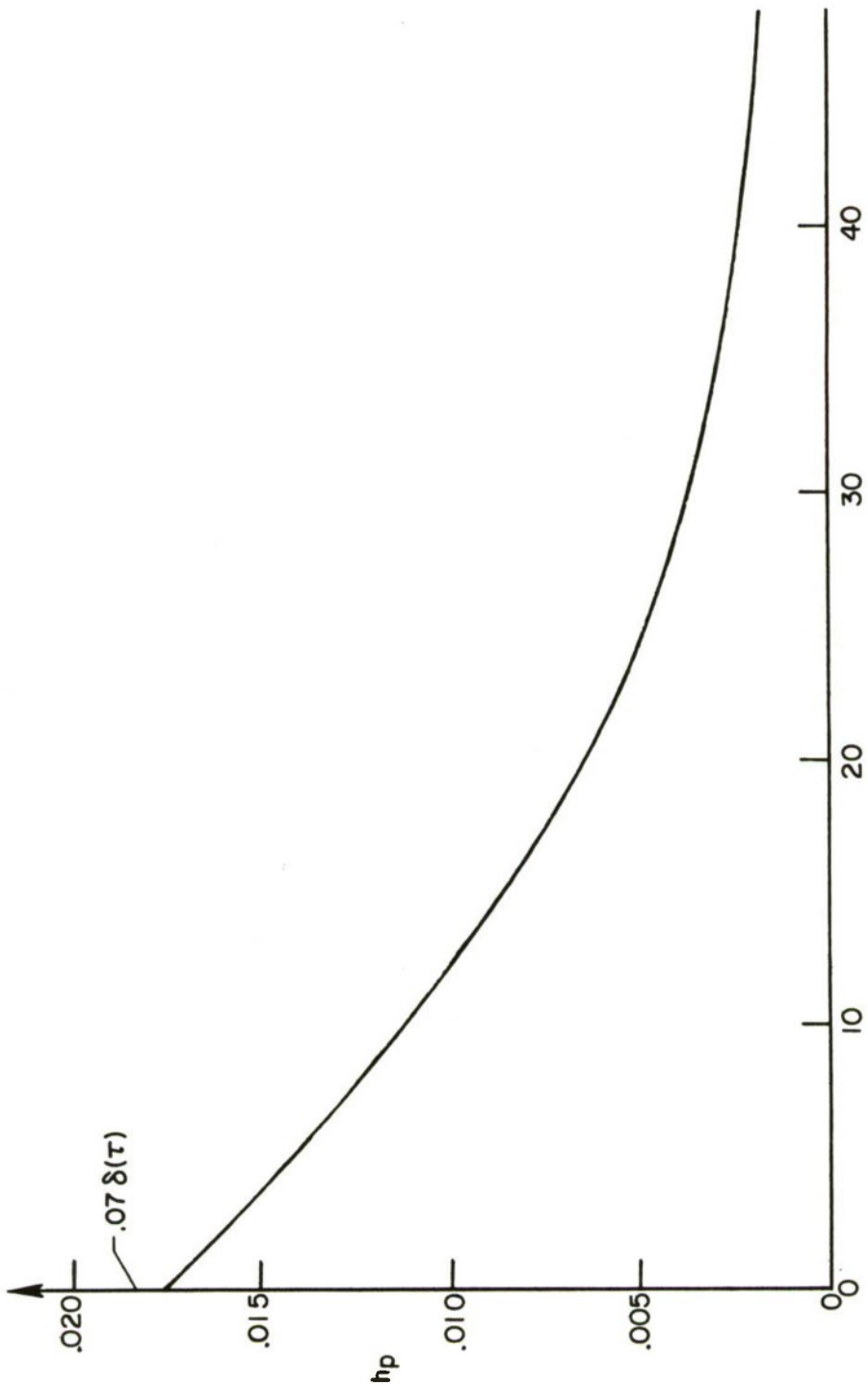


Figure 9.- Impulse response for pilot model used to study truncation effects.

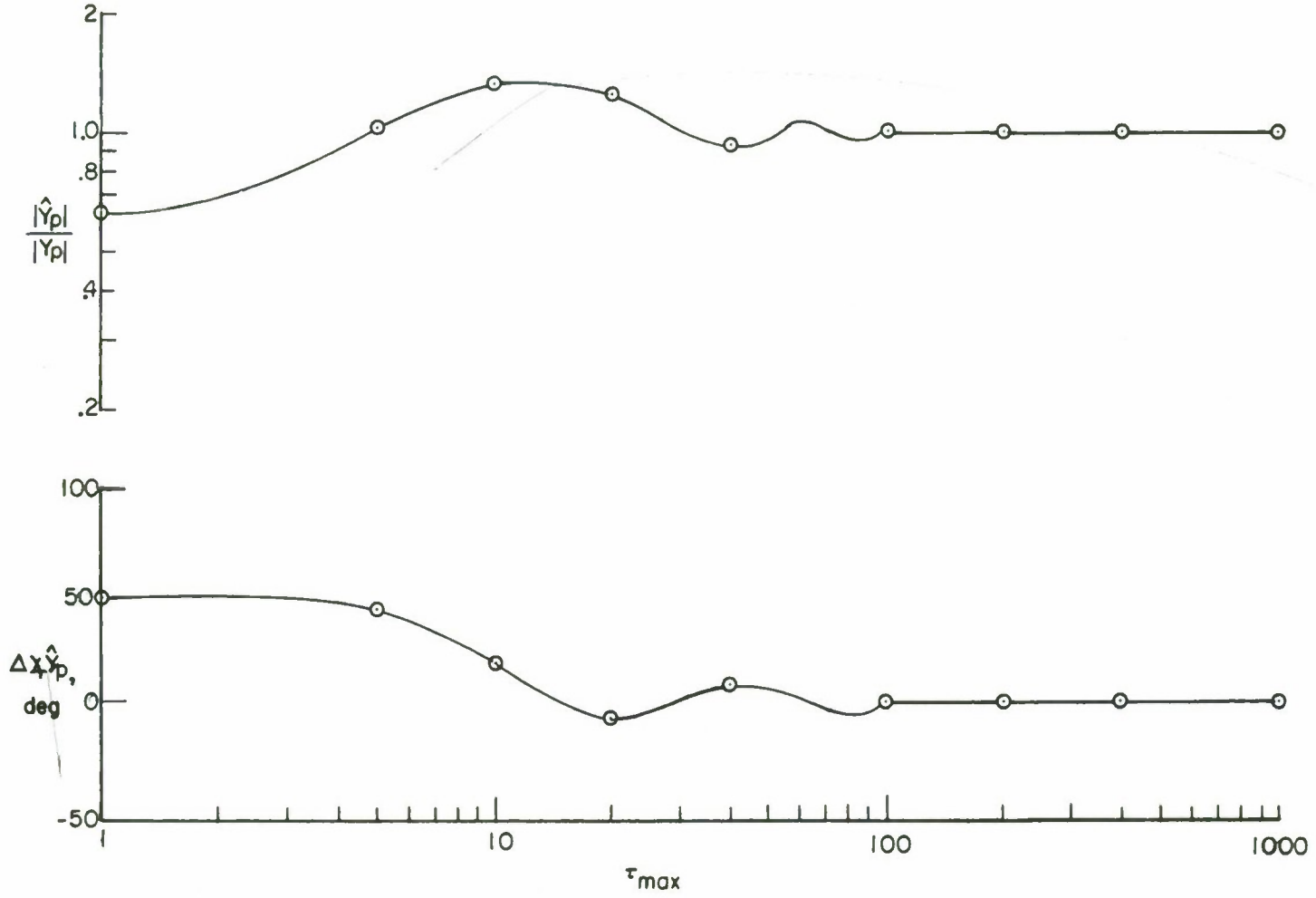


Figure 10.- Effect of memory time on Y_p. $\omega = 0.157$ rad/sec.

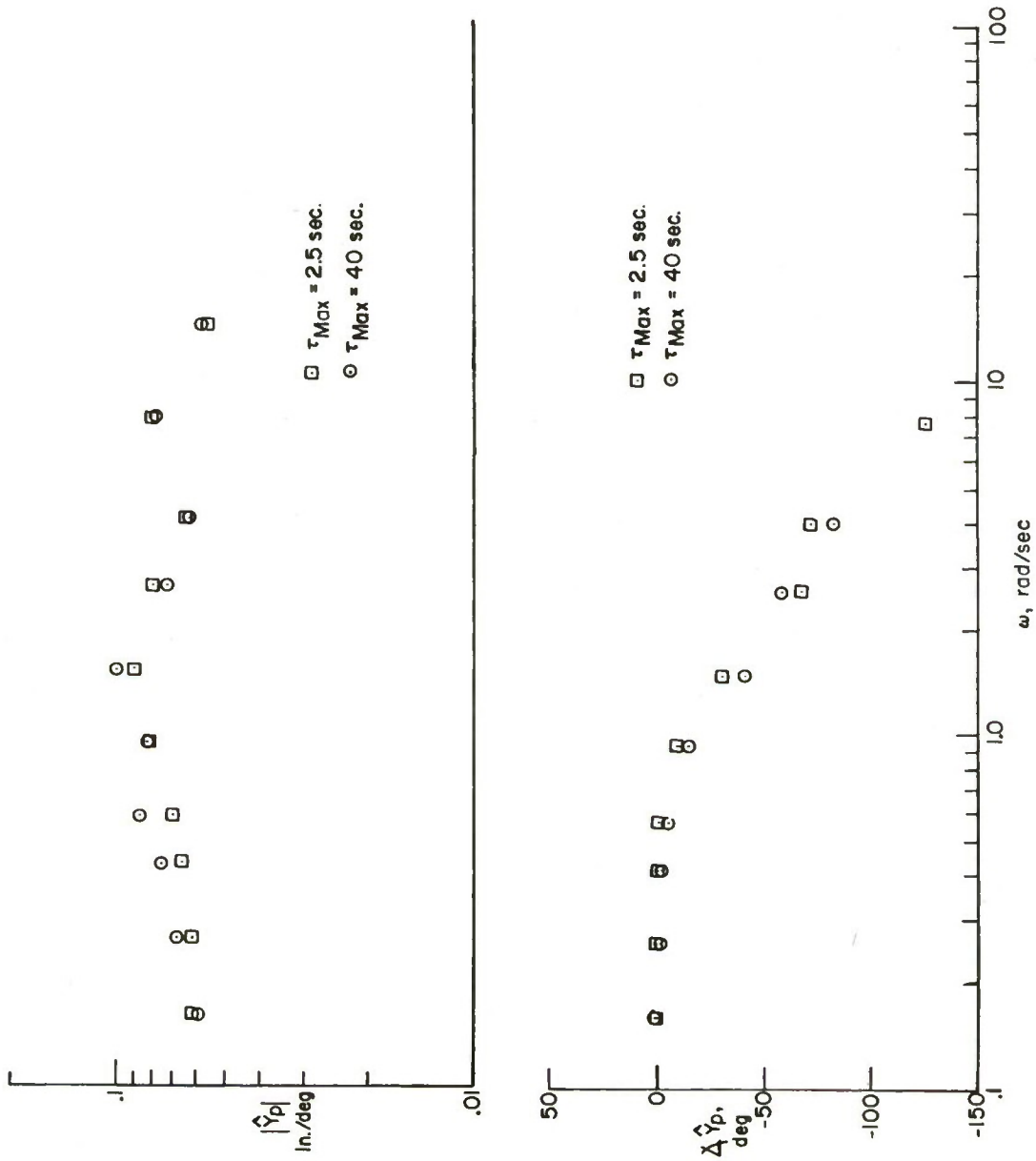


Figure 11.- Comparison of time domain results.