DEPARTMENT
OF
COMMERCE

MISCELLANEOUS
PUBLICATIONS
OF THE
NATIONAL
BUREAU
OF
STANDARDS

NOS. 265, 267 - 270

UNITED STATES DEPARTMENT OF COMMERCE • John T. Connor, *Secretary*
NATIONAL BUREAU OF STANDARDS • A. V. Astin, *Director*

Systems Engineering in Ceramics

Proceedings of a Symposium
April 19, 1964

Held under the auspices of the Ceramic Educational Council and National Institute of Ceramic Engineers of the American Ceramic Society, with the cooperation of the National Bureau of Standards, and under the sponsorship of the Office of Naval Research. The symposium took place at the 66th Annual Meeting of the American Ceramic Society in Chicago, Illinois.

National Bureau of Standards Miscellaneous Publication 267

Issued May 1, 1965

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 20402
Price $2.00

Library of Congress Catalog Card Number: 65-60036

Contents

Page

Introduction iv

1. Systems Engineering and Modern Engineering Design 1

Allen B. Rosenstein
University of California
Los Angeles, California

2. Basic Concepts of Systems Engineering 11

H. E. Koenig and C. M. Bacon
Dept. of Electrical Engineering
Michigan State University
East Lansing, Michigan

3. Process Control in the Ceramic Industry 39

Ernest Hucke
Honeywell Industrial Products Group
Philadelphia, Pennsylvania

4. Application of Computers in Design and Control of Systems 81

Joseph L. Hammond, Jr.
School of Electrical Engineering
Georgia Institute of Technology
Atlanta, Georgia

5. Statistical Methods in Systems Design 109

W. C. Hackler
North Carolina State of the
University of North Carolina
Raleigh, North Carolina

6. Systems of Interest to Ceramic Engineers 119

A. H. Hurst and E. R. Bullock
Instrument Dept.
Minnesota Mining & Manufacturing Co.
Hawthorne, California

Introduction

The last decade has witnessed the increased interest of Ceramic Engineers in the development of a number of engineering techniques for the analysis and synthesis of groups of interconnected elements or components. These techniques, which have been developed by many engineering disciplines, have been found to be applicable and useful for studying overall system behavior independently of the disciplinary origin of the components. It is not surprising that an emerging area called "Systems Engineering" has begun to encompass these techniques and methods of attack, and to emphasize their usefulness when attention is focussed on the interaction of two or more elements.

Although the average Ceramic Engineer is unlikely to shift his professional activities wholly into this new area, Systems Engineering overlaps the interest of Ceramic Engineers in several important areas. Where this overlap occurs, Ceramic Engineers can benefit from well established techniques and methods. The design and use of process control systems is possibly the most important single area where such overlap takes place. Other areas of overlap include the design of manufacturing operations, pilot plants, and systems for experimental testing and evaluation.

The Symposium on Systems Engineering was designed to present an overall concept of this emerging discipline with emphasis on topics of interest to Ceramic Engineers. In Chapter 1 Dr. Rosenstein introduces the concept of systems engineering and discusses its boundaries and its relation to engineering design in other disciplines. In Chapter 2 Drs. Koenig and Bacon present a general discussion of some of the basic concepts used in any systems engineering problem.

Chapter 3, by Mr. Hucke, presents a discussion of the important class of process control systems. In the next two chapters, Drs. Hammond and Hackler consider two of the more important system engineering techniques, namely: computer methods and statistical methods. The final chapter, by Mr. Hurst and Mr. Bullock, ties together the material of the first five chapters by considering in detail several examples of systems of interest to Ceramic Engineers.

W. E. Moody, Chairman
Georgia Institute of Technology

H. F. McMurdie, Editor
National Bureau of Standards

W. E. Coffeen
M & T Chemicals, Inc.

Winston Duckworth
Battelle Memorial Institute

J. L. Hammond
Georgia Institute of Technology

E. C. Henry
General Electric Co.

G. E. Lorey
Missouri School of Mines

G. C. Robinson
Clemson College

Systems Engineering and Modern Engineering Design

Allen B. Rosenstein

1. Introduction

Within the past ten years there have appeared with increasing frequency books, articles, conferences, and monograms dealing with system engineering, system analysis, system design, the systems approach, the design of systems, system theory, and problems of systems engineering.[1,2,3,4,5,6,7,10] The number of publications and the stature of their authors does not allow the dismissal of the subject as a passing fad. The breadth of engineering activity involved in even a cursory examination of recent publications is of interest. Articles involving Systems Engineering have appeared in Aeronautical Engineering, Electrical Engineering, Product Engineering, Industrial Engineering, Oil and Gas, Instrument and Control Systems, Control Engineering, Industrial Chemistry, I.S.A. Jr., Coal Age, Modern Materials Handling, and Machine Design. It is therefore obvious that whatever Systems Engineering may or may not be, it is non-sectarian and encompasses activities that are of concern in all phases of engineering. On the other hand undergraduate college offerings akin to Systems Engineering are rather limited and even graduate programs are not extensive.

This article will attempt to give some perspective to the emerging phenomena called Systems Engineering. An examination of the technological needs of our society should provide some understanding of the almost simultaneous appearance of the subject on so many fronts. Once we have satisfied ourselves as to the authenticity of its antecedents, it will be useful to examine a brief history of the subject. With needs and history given, we will then have a go at establishing a definition or definitions of Systems Engineering. Since the definitions usually do not suffice, we will attempt to illustrate the system concept with descriptions of System Engineering activities and elements.

The question of system design versus component design cannot be avoided and will require appropriate space. Systems design and component design cause us to consider the relation of design in general to the practice of engineering. This in turn calls for careful examination of the concept of Modern Engineering Design. The modern description of the design process will take us back to our original subject for it needs careful comparison with the current exposition of Systems Engineering. Side comments will be in order to orient the role of Operations Research.

It has been observed that an invention or an idea will often appear spontaneously and independently in many places. In retrospect this has been found to result from the convergence of new or growing needs of society with the appropriate ideas, bits and pieces of technology and new insights into old problems. The pressure to generate the ideas and methods attributed to Systems Engineering stems directly from the needs of 20th century society. As our frontiers have disappeared, man has turned to technology to furnish the "good life" in a rapidly shrinking, crowded world. Our interdependence upon one another has increased in direct proportion to the population increase. The race to maintain or improve the operating efficiency of society has required that the systems and mechanisms that serve the society also become increasingly complex and interdependent.

Goode and Machal have provided statistics to illustrate the above.[4] They note that the world population increased from 800 million in 1750, to 1200 million in 1850, and 2400 million in 1950. Maximum transportation speeds went from 40 mph in 1850, and 100 mph in 1900, to commercial transport speed of 350 mph in 1950 and supersonic transport planes of over 1200 mph in the 1960's. Our communication systems are a good indication of increasing complexity. U.S. telephones jumped from 350,000 in 1900, to 55 million in 1955.

The return on investment in machinery has increased productivity with the attendant opportunity for more leisure time and more education. As a good example of this, one farm worker supported six people in 1900, and fifteen in 1950. The penalty has been increased complexity, more sophisticated and difficult design problems, and greater interdependence. The telephone industry in 80 years has gone from simple manual switchboards with direct connection to world wide automatic switching wherein machines select the "best" of alternate routes, place the call, record the charges, and mail you the bill.

2. History

Hall of the Bell Laboratories has attempted to trace the historical development of Systems Engineering.[1] R.C.A. is credited with recognizing a need for a systems approach in the 1930's for the development of a television broadcasting service. World War II witnessed a quantum jump in the complexity of systems which were met with tools and techniques ranging from Operations Research to more sophisticated control theory. Although the Bell Telephone Laboratories Incorporated is credited with using the term "Systems Engineering" in the early 1940's, widespread usage of the terms did not appear until the 1950's.

Hall points out that, "It is hard to say whether increasing complexity is the cause or effect of man's effort to cope with his environment. In either case a central feature of the trend has been the development of large and very complex systems which tie together modern society. These systems include abstract or non-physical systems, such as government and the economic system. They also include large physical systems like pipe line and power distribution systems, transportation and electrical communication systems." In either event the results have been very clear. The need for communication systems, military systems, transportation systems, and managerial systems of increasing complexity has called for the development of models, analytical techniques, and design methods of increasing sophistication and power. In the past decade this call has been met with increasing success. The knowledge and experience gained has opened still more doors. Process control systems, for example, can be analyzed in greater detail and regulated with increasing precision. Our insight into the modeling, organization and mechanization of man-made systems has now prepared us for a massive assault upon the secrets of the biological systems.

One important aspect of today's technology that is often overlooked lies in the increasing importance of engineering system decisions. Because of the size of our systems and the extent of their interaction with society, the opportunities for catastrophic (or near catastrophic) engineering failures are increasing. While engineers still are called upon to design simple two lane highways, some engineers are expected to design complete transportation systems. These systems in turn will control the growth of entire regions for many generations. The author is particularly sensitive on this point since he lives in a city whose "solution" to the mass transportation problem has created an equally serious hazard called smog. The total dependence of our national military position upon approximately one dozen highly complex weapons systems is a further example which should give us cause for concern.

3. "Definition" of Systems Engineering

As a term becomes popular it is customary to seek a concise definition which clearly illustrates its meaning. If the term is applicable to a process - and a very general process at that - the lucid definition more often than not becomes elusive. For a unique definition of such processes, it usually becomes necessary to look to a description of the elements which comprise the process. On the other hand, there have been those useful activities that have resisted precise categorization in their formative years, but which have grown up into a very respectable, well-defined middle age.

It is too early to make categoric statements about the ultimate acceptance and usage of the term System Engineering. There is not yet a consensus of opinion about the activites encompassed. However, large areas of agreement are beginning to appear as each author in turn takes a crack at defining the boundaries of the subject.

Since this author has developed his own views on the ultimate role of Systems Engineering, it seems only fair to report the opinions of others. A recent panel discussion of System Theory[2] produced the following:

"System Theory is concerned with the general problems of modeling and design of systems."

"System Theory may be thought of as mathematics applied to engineering in the same sense that the study of devices is physics applied to engineering problems."

"As devices exploit physical phenomena in solids, fluids, gases and plasmas, so systems exploit relationships and transformations to realize their purposes."

"System Theory is the abstraction, generalization and successor to circuit, communication, control, and computer theory."

The above statements concern System Theory. J. A. Morton of the Bell Telephone Laboratories describes the Systems Engineering Method as follows:[3]

"The Systems Engineering method recognizes each system is an integrated whole even though composed of devices, specialized structures and sub-functions. It further recognizes that any system has a number of objectives and that the balance between them may differ widely from system to system. The methods seek to optimize the overall system function according to the weighted objectives and to achieve maximum capability of its parts."

Goode and Machal in their book, "System Engineering, An Introduction to the Design of Large Scale Systems,"[4] have the following to say:

"The lack of formal definition does not prevent us from noting the characteristics which are frequently present in large scale systems. Each such system has a certain integrity. It may or may not be rigidly controlled from some central point, but in every case, all the parts of the system have some common purpose; in some sense, they all contribute to the production of a single set of optimum outputs from the given set of inputs, with respect to some appropriate measure of effectiveness."

The current interest in Systems Engineering is not confined to this country. In England, Gosling has produced a stimulating book entitled, "The Design of Engineering Systems."[5] Gosling describes systems as follows:

"Some engineering artifacts are most easily analyzed, described or designed as an assembly of simpler parts. Artifacts of this kind are called <u>systems</u>. Some systems have the property that flowing through them are streams of some 'working fluid' (which may be matter, energy, or information), in such a way that the 'working fluid' passes in turn through many parts of the system, which is in consequence termed a sequential (or flow) system. Examples are a chemical plant, an electrical power distribution network, a digital computer, a sewer system. Systems which do not have this property are termed associative systems of which examples are a motor car, an aircraft, or a bridge - - it is with (sequential) systems that the theory of system design has primarily been developed."

If the concept of flux is accepted for flow and a force flux is considered to be pervoying through a structure, then even a static structure may be reconsidered in terms of Gosling's definition of a sequential system.

4. The Content of Systems Engineering

"Engineering" is the name of a profession, and at the same time the description of the activity in which engineers engage. The foregoing definitions of Systems Engineering can be given more body by examining the important ingredients of that activity. Examination of three of the more penetrating studies of the subject reveal a remarkable unanimity as to what constitutes the body of Systems Engineering.[2,3,5] With apologies to the respective authors, if their works have been misinterpreted, summaries of the particular parts of their publications in which they have listed the essential elements of Systems Engineering follows:'

Phases of Systems Engineering - Arthur D. Hall

I. Systems Studies:
   1. To investigate the Environment
   2. To create background of Information

II. Exploratory Planning
   1. Problem Definition - data gathering, customer requirements, inputs, outputs
   2. Objective selection - search for alternatives, criteria for optimum
   3. Syntheses of system - alternative systems to satisfy objective with data for decision.
   4. Analyses of system - consequences of alternative systems
   5. Selecting of the Best system
   6. Communicating results

III. Development Planning
   1. Recycling of above in more detail
   2. Experimentation may be conducted

IV. Studies During Development
   1. Development and manufacture of first models
   2. Final evaluation of system in its working environment

V. Current Engineering
   1. Continues as long as system is used
   2. Feedback to develop "know how" for future systems

Approach to Systems Engineering Problems - Harold Chesnut[6]

I. Problem Formulation - customer requirements or needs; objectives; restraints including inputs, environment, human factors; weighting functions.

II. Synthesize System - alternate solutions

III. Way to Make System Synthesized

IV. Measure and Compare with Objectives

V. Refine

The Design of Engineering Systems - W. Gosling

I. Completion of Specifications - Inputs, Outputs, Environment, Measures of Value

II. Search for Alternative Solutions

III. Feasibility Check - Physical laws, Environment Compatibility including human laws and customs, cost, company capability

IV. Modeling

V. Reliability Study

VI. Selection of a Best System

VII. Evaluation by Trial

VIII. Collection of system performance statistics

The ideas embodied in these operational statements of Systems Engineering are quite familiar to engineers. What does perhaps appear new or perhaps only reemphasized on a new scale are certain clearly stated concerns. Three of these that stand out are:

1. A concern with the overall problem ranging from the initial conception to the recording of the operating life performance.

2. A concern with optimization of system performance in terms of formal measures of value.

3. A concern with the needs of the customer that are determined on several levels.

In addition to the structure for Engineering System studies, a number of tools of especial value to the Systems Engineer have been given. High on most lists are probability, statistics, logic, and computing. Often included are queueing theory, game theory, linear programming, dynamic programming, simulation, information theory, automatic control theory, topology, philosophy, economics, psychology and language in addition to the more traditional topics of engineering, mathematics and science.

The list contains enough items that have come from the field of Operations Research to warrant a comparison. Since the Operations Research proponents have had even greater difficulty in defining their bailiwick than the Systems Engineers, it is not surprising that a ready distinction is not available. The writings of O.R. experts do not necessarily aid us. Churchman et al.[7] list the following phases of operations research which compare rather closely with those of Systems Engineering.

1. Formulating the problem
2. Constructing a mathematical model
3. Deriving a solution from the model
4. Testing the model and the solutions derived from it
5. Establishing controls over the solution
6. Putting the solution to work: Implementation

Hall believes that Operations Research is usually concerned with the operation of an existing system while System Engineering emphasizes the planning and design of new systems. Goode and Machal state that the operations research is primarily interested in making procedural changes while the systems engineer is primarily interested in making equipment changes. The line may be seen although it is drawn rather thin.

5. Component Engineering

Those readers who have been patient enough to remain with us this long have certainly begun to ask themselves, "When is a component not a component, but a system, and who decides when to apply the new Systems Engineering in preference to the old components engineering?" An air transportation complex is complicated enough to warrant the designation of engineering system. The jet transport would be a component in the transportation system, and yet its design would exercise all of the elements we have attributed to Engineering Systems. The aeroplane's radar and communication electronics are components of the craft but are substantial systems in their own right. We can continue down to a radar component such as a pulse transformer. The transformer in turn possesses a magnetic circuit, a thermal circuit, a mechanical structure, electric circuit, and insulating system all of which are functions of their geometries and materials.

It is generally agreed that the system-component distinction is an individual decision to be made from the point of view of the engineer who is faced with the decision. While the system-component contrast is admittedly made on an apparently arbitrary basis, the distinction in the final analysis will be found to be one of engineering survival. As the number of individual items in a project increase, the amount of information that can be economically processed per item must decrease. When the description of some items have been reduced to either a statistic or hardly more than an input-output relationship, we are inclined to classify these items as system components.

The procedures, methods, models, and tools of system design and of component design are essentially the same. The items and information content of complex, large scale systems have spurred the development of models and techniques for dealing with what may be called high density traffic in matter, energy, and/or information. But as is usually the case, these techniques are finding increasing application in the design of components which in turn have become more sophisticated.

6. Modern Engineering Design

As the differences between Systems Engineering, Component Engineering, and even Operations Research become increasingly difficult to discern, we should re-examine our ordering of engineering and of engineering design. Such a study was begun several years ago by the U.C.L.A. Department of Engineering. A committee spent several years exploring the role of engineering, the content of modern design, and its place in the engineering of the future.

The committee concluded that design taken in its broadest sense is the essential activity of engineering. Design was defined as an iterative decision making process to develop systems to optimize the value of resources. In this context, design is intended to encompass the entire process from the appreciation of a need of society to the ultimate disposal of the systems created to satisfy the need. The research required to provide knowledge necessary to implement the design is considered an element of the process.

Studies of the activities of prominent engineers convinced the committee that the design process could be described (and therefore taught). A horizontal and vertical structure was observed to exist in the design process. The vertical structure which has been called the "Morphology of Design" gives the complete step by step design procedure.[9] The horizontal structure consists of those elements and ingredients of the process which are repeated over and over again throughout the design. These are collected into what has been called the "Anatomy of Design". The elements of the Design Anatomy, listed below are the essential ingredients of the design process which are applicable to all design, whether system or component. A knowledge of these elements, their interrelationships, and the techniques associated with each element provides a foundation for all modern engineering design.

ANATOMY OF DESIGN

1. IDENTIFICATION OF THE NEED.

   As analysis of whether or not a need exists is required. Often a need, as originally stated, is not what is really needed to satisfy the final objective.

2. INFORMATION COLLECTION AND ORGANIZATION.

   All factors which relate to the system need to be considered. Where necessary, experiments must be devised to obtain data otherwise unavailable.

3. IDENTIFICATION, STATEMENT, AND MODELING OF SYSTEM VARIABLES.

   All factors influencing the system - the so-called "boundary conditions".

   Engineering systems, subsystems and components can be analyzed into basic elements which, when described or prescribed in appropriate detail and when properly synthesized, will constitute the design of the system.

   a. INPUTS: Those resources and other environmental factors which are converted (or modified) by the system in question.

   b. OUTPUTS: That which is produced by the system.

   c. TRANSFORMING MEANS: Used to obtain the relationship between inputs and outputs.

   d. CONSTRAINTS: All elements and factors which express limitation and/or need to be accounted for in the design.

4. CRITERIA DEVELOPMENT FOR OPTIMUM DESIGN. The rules for judging relative merit.

   a. Development of Value Systems

   b. Criteria relationship among values

5. SYNTHESIS AND ANALYSIS.
   Evolving of systems to convert the inputs into the desired outputs. At this step only the requirement of realizability is usually met. Analysis forms the feedback loop around Synthesis.

   a. Physical Realizability
   b. Economic Worthwhileness (realizability)
   c. Financial Feasibility (realizability)
   d. Producibility
   e. Maintain ability
   f. Reliability

6. TEST EVALUATION, AND PREDICTION.

7. OPTIMIZING.
   Optimizing (Maximizing the performance. Reduction to "best" system with available knowledge.)

8. DECISION.

9. ITERATION.

It is recognized that the above operations are found throughout the design process. Many iterations will be taken around several or all of them. In particular the engineer continually re-examines his previous findings and decisions in the light of new information.

10. COMMUNICATION, IMPLEMENTATION, AND PRESENTATION.

The similarities between the content of Systems Engineering and the Modern Design Process are too great to be ignored. From the pedagogical standpoint an organized view of the general design process and its important constituents offer an excellent opportunity to provide the student with insight into and capability for undertaking the essential functions of engineering regardless of whether he is employed at a systems or component level.

7. Conclusion

Labels are, of course, not too important. In the long run it does not matter whether or not the titles Systems Engineering or Component Engineering become fully accepted general terms. It is, however, important that their practitioners are prepared to exploit to the utmost all of the tools and experience that modern engineering can provide. Many, if not most of the present day significant engineering undertakings have become too large for a single individual to encompass. Divisions of responsibility and labor cannot be avoided. Yet regardless of how the distinctions are made, there is a greater need than ever for a full appreciation of the functions and responsibilities of all facets of the design process.

I would in particular like to stress the work "responsibility". No matter whether we are working at a system level or a component level, every engineering project is part of some larger organism or system. It consequently becomes a subsystem to this larger system. All of us therefore in this sense, design components which ultimately serve a greater system. As professional engineers there exists a responsibility to continually question and re-examine the inputs to our subsystem, the weighting functions we use to optimize value, and to check the ultimate performance of our products. There are the individuals who are willing to blindly accept specifications, perform their computations, and submit their designs to other departments for implementation with little further interraction. Indeed, large engineering projects tend to organize themselves along these lines. However, the man who does not participate in any of the value judgments is not functioning as a professional engineer. He is serving as a technician, a technologist or perhaps as a human computer. In any event, it is a subprofessional activity that possesses a real possibility for future relegation to machine operation.

Within every large engineering organization there is an increasing need for specialists in breadth as well as the continuing need for the specialists in depth. It is, of course, this need for specialization in breadth that has forced the development of "Systems Engineering". Yet we have seen that the Design Process is the same for both the specialists in depth and in breadth, the systems and the components engineer. The basic tools, models, techniques are also very much the same. We need only to concede that at this time that the specialist in breadth may have a greater need for traffic handling techniques. Since the differences in the specialists do not appear in the general design process, or in their procedures or models, we have to look further for significant variances. The most important will be the respective storehouses of the characteristics of the real world, that each will accumulate to perform their given tasks. If we can allow ourselves the luxury of a rather broad generalization, we could observe that the specialist in depth will orient toward microscopic phenomena while the specialist in breadth will be concerned with the macroscopic.

The clearest distinction lies in the kind of information processed rather than the quality. If we assume that men of equal intelligence are capable of processing information at equal rates, then the specialist in breadth dealing with a large number of items is forced to reduce the amount of information associated with each item. Conversely, the specialist in depth who restricts the number of items involved in his designs, will be free to increase the degrees of freedom allowed each item with an attendant increase in information per item. To this end, it is interesting to note that the complexity of important organic molecules often exceeds that of our most intricate processing or communcation systems.

References:

1. "A Methodology for Systems Engineering," Arthur D. Hall, Van Nostrand 1962.

2. "What is System Theory and Where Is It Going?", A Panel Discussion, IEEE Trans. on Circuit Theory, June 1963.

3. "Integrating of Systems Engineering with Component Development", J. A. Morton, Electrical Manufacturing, August 1959.

4. "System Engineering, An Introduction to the Design of Large-Scale Systems", Harry H. Goode, Robert E. Machal, McGraw-Hill, 1957.

5. "The Design of Engineering Systems," W. Gosling, Wiley, 1962.

6. "Preliminary Survey of Systems Engineering Problems", H. Chesnut, General Electric Laboratory Report No. 60 GL 5, January, 1960.

7. "Introduction to Operations Research", C. W. Churchman, R. L. Ackoff, Wiley, 1957.

8. "Design as a Basis for a Unified Engineering Curriculum", UCLA 1962, A. B. Rosenstein and J. M. English.

9. "Introduction to Design", Morris Asimow, Prentice-Hall, 1961.

10. "Operations Research and Systems Engineering", C. D. Flagle, W. H. Huggins, R. H. Roy, Johns Hopkins Press 1960.

Basic Concepts of Systems Engineering

Herman E. Koenig and Charles M. Bacon

1. Introduction

The appeal of system theory to scientific workers in many disciplines has increased tremendously in the last decade. The work "system" is currently applied in a host of scientific endeavors including many fields of study not directly related to the Physical Sciences and Engineering, such as Economics, Sociology, Biology, Transportation and Supply, etc.

In an effort to apply "systems concepts" to their own particular problems, workers in these latter areas have turned to the engineer and mathematician for a clear concise exposition of system theory. This responsibility has caused the engineer to re-examine the entire area of systems analysis, with a view toward defining the discipline more clearly, generalizing the concepts and broadening the base of application.

The authors of this paper assume that the primary application of system theory in ceramic engineering involves feedback control problems, as they arise in materials handling and process control. If a satisfactory process control system is already in operation, system theory has little to offer in the way of providing new and useful information. However, if a process control system is to be introduced, improved or replaced, the performance characteristics of the proposed system can be completely analyzed to see if it satisfies specifications. If it does, the process engineer can be confident that the system, when built, will meet those specifications within an acceptable tolerance level. Such a procedure is called <u>system analysis</u>. If on the other hand, the analysis shows that a proposed system does not meet specifications, then system theory is used to alter certain components of the system so that the altered version does meet specifications. This procedure is called <u>system design</u>.

All quantitative system analysis and design procedures are based on what is called a <u>mathematical model</u> of the system. Mathematical models of a given system may vary widely depending on the performance characteristics to be modeled, the accuracy to be realized and the analytical techniques to be used in solving the model. Prior to the second World War, quantitative system theory existed only in a crude form; being concerned largely with steady-state performance characteristics. During the war, an intense research effort, centered about the dynamic analysis and design of radar and gun control systems, brought forth transform methods as an efficient means for analyzing systems characterized by linear differential and difference equations. The transfer function became, and still remains to a large extent, the standard mathematical model for linear systems. Many excellent texts exist which deal with system analysis and design using the transfer function approach.

Since about 1958, the profound need for more effective methods for analyzing the nonlinear systems encountered in missile and launch vehicle systems, coupled with the availability of computing machines, has shifted the emphasis in system theory from transform methods to what are called <u>state-space</u> models. These models are ideally suited for system simulation on either analogue or digital computers and serve as the basis for modern system theory.

In a tutorial development such as this, space does not permit the development of a working knowledge of modern system theory. Entire books are devoted to this subject. This paper, rather, attempts to develop, mostly through the use of examples, an appreciation for the main stream of what modern system theory is about so that the reader might evaluate its potential applications to the process control systems involved in ceramic engineering. To this end, two aspects of system theory are delineated; 1) the problem of developing mathematical models of the system and 2) the "processing" or solution of the model to ascertain the performance characteristics of the system. The first aspect is sometimes referred to as "system simulation" and in the case of complex process control systems containing mechanical, hydraulic and electrical equipments, frequently represents the most difficult aspect of system analysis. Fortunately, systematic modeling procedures have recently been defined. A discussion of these procedures is presented in Section 3 after developing in Section 2 some of the prerequisite system concepts pertaining to measurements and component models. Section 4 considers several methods of solution of the system model as well as the derivation of the system transfer function. Finally, the important concepts of feedback and stability are illustrated by example in Section 5.

2. Variables, Measurements and Component Models

A system is the sum of its parts. The performance characteristics of a system are, therefore, a function of the performance characteristics of the individual components or parts which comprise the system. It follows that a mathematical model which describes the characteristics of a system, can only be obtained when the mathematical models of the individual components are available. And yet while the models of the individual components are necessary to the formulation of the system model, such information is insufficient. Three given components interconnected in two alternate configurations may give entirely different system performance. Thus, in general, two pieces of information are required before a mathematical model of a system can be derived; 1) a model of the performance characteristics of the individual components and 2) a mathematical statement or model of their interconnection.

It is implied that the mathematical model of each component is characteristic of the component itself and independent of the way in which the component is connected with other components in the system. That is, like the name plate and serial number, the mathematical model of the component characteristics goes with the component, to all "customers".

To illustrate this point and show more explicitly what is meant by a model of a component, consider the hydraulic lift mechanism in Fig. 2.1. To develop a mathematical model of this system, we first attempt to separate the system into components or subassemblies for which mathematical models can be more easily obtained. The system has been divided into three subassemblies as indicated by the dashed lines. These subassemblies can be regarded as "black boxes" interconnected as shown in Fig. 2.2, where the letters A, B, H. and J indicate points of union between the components --- points at which the performance characteristics of each component as a separate entity are to be measured and characterized mathematically.

Assume that only vertical forces and displacements with respect to the fixed reference G are of interest. The control level assembly merely translates a given force and displacement "input" at point A into another force and displacement at point B. At point B, the hydraulic subassemb y receives this mechanical input and translates it into a much larger force and displacement at point H. The level G-H-J transmits this force and displacement to the mass at J. Each subassembly model must relate, mathematically, the mechanical quantities of force and displacement at the terminals where connections are made to other components. To develop these models each subassembly can be considered in greater detail.

The hydraulic subassembly, for example, is divided into the four basic components shown in Figure 2.3. It is assumed that the characteristics of the hydraulic pump in Fig. 2.1 and the fittings between the pump and the control valve can be included in the mathematical model of the control valve. From Fig. 2.3, it is evident that the control valve translates the mechanical force and displacement at point B into hydraulic fluid flow and pressure at points C and D. Each hydraulic line exhibits a pressure differential over its length which is a function of the fluid flow through the line. Finally, the hydraulic ram translates a difference in pressure and flow rate at points E and F into mechanical force and displacement at point H. Thus, the hydraulic subassembly is considered as a system having four components each having "terminal" characteristics that can be modeled mathematically. From the models of these components we can hopefully derive a model of the hydraulic subassembly, then a model of the entire system.

Note that exactly _two_ physical quantities or variables have been cited in describing the effect of one component on another --- mechanical force and displacement or pressure differential and fluid flow. In the ensuing development, it is assumed that, in general, _two_ such complementary variables are required to describe the effect of one component on another at their common interface. The mathematical model of the control valve, for example, is complete when the complementary pair of variables at the "output" have been expressed as a function of the complementary pair of variables at the "input". This functional dependence can be established by removing the component from the system and subjecting it to certain performance tests. During such performance tests, the vertical displacement at point B in Fig. 2.3 must be measured. But this measurement must be taken with respect to an appropriately selected reference point. If point G is used as a fixed reference, the vertical displacement of B with respect to G at any instant of time can be described by a real number. Similarly a pressure measurement at point C has no meaning unless a reference is specified. The selection of point D as reference implies a pressure differential between points C and D on the valve. If gage pressure at C is desired, the local atmospheric pressure is taken as reference. Thus, a well defined reference is required. A similar statement can be made about the measurement of other physical quantities such as voltage, rotational displacement and temperature difference. This reference requirement _implies_ a "two-point" property of physical measurements.

Component mathematical models are conveniently classified with respect to the number of points (or terminals) on the component at which physical measurements are made. A model of the control valve in Fig. 2.3 is designated as a four-terminal model when it is based on measurements between points B and G and between points C and D. Similarly, the hydraulic line CE can be modeled as a two-terminal component by considering measurements between points C and E. Fig. 2.4a shows schematically a _two-terminal_ physical device subjected to a performance test. A "driver" excites the component in order to simulate its behavior in a system while two complementary physical quantities (pressure difference and flow rate, for example) are measured with respect to terminals A and B. One of these quantities exhibits a "_difference_" property with respect to A and B, i.e., it assumes different values at A and B. For this reason, it is identified as the _across_ variable $x(t)$; in general, a time-varying quantity. The second quantity exhibits a "_common_" property with respect to A and B. That is, its values at A and B are equal. Thus, it is identified as the _through_ variable $y(t)$; also time varying. The meters are to be regarded as _conceptual_ since the component may be electrical, hydraulic or mechanical and the physical quantities to be measured will vary accordingly. The polarity signs on the meters indicate that they furnish both positive and negative values. Thus, a meter indication of opposite sign is obtained if the meter connections are momentarily reversed.

The mathematical model of the component in Fig. 2.4a describes _two_ distinct features of the conceptual measurement apparatus. First, the orientation of the meters and their manner of connection to the component are represented by the oriented line segment of Fig. 2.4b. Second, the time variation of one measurement, as a function of the second measurement is approximated by an equation; one possible form appearing in Fig. 2.4c. The line segment, hereafter called the _terminal graph_, along with the equation, hereafter called the _terminal equation_, constitutes a mathematical model of the two terminal component.

If, for example, the component being subjected to performance test is the length of hydraulic line in Fig. 2.5a, the through and across variables become hydraulic fluid flow and pressure difference. The conceptual measurement apparatus of Fig. 2.4a is represented in the mathematical model by identifying the ends A and B of the hose by the pair of points a and b on the terminal graph shown in Fig. 2.5b. The orientation of the graph signifies that $p(t)$ represents the pressure differential of _a with respect to b_, and that $g(t)$ represents positive fluid flow _into terminal a and out of b_. If the orientation of the terminal graph is reversed, then _both_ pressure and flow measurements are reversed. A typical measured static pressure-flow characteristic appears in Fig. 2.5c together with a terminal equation of a straight-line approximation of the recorded characteristic. The model is incomplete unless both the terminal graph and terminal equation are given. One is useless without the other!

The terminal graph alone tells nothing of the functional relation between p(t) and $\dot{g}(t)$ and the terminal equation alone furnishes no information about the measurements represented by p(t) and $\dot{g}(t)$.

In each mathematical model, <u>one</u> across variable and <u>one</u> through variable are used to describe the performance characteristics of each type of physical device (electrical, mechanical, hydraulic, etc.). Table 2.6 lists some of the important physical quantities and their classification as through or across variables. Many practical devices can be modeled on the basis of only two measurements related by a single terminal equation and are thus classified as two-terminal components. Table 2.7 lists the most common two-terminal components together with their terminal equations.

Depending upon the approximation which can be tolerated in the analysis of a system, terminal equations can range in complexity from simple linear equations to nonlinear differential equations. The analysis of a system may become exceedingly difficult if it is necessary to use nonlinear terminal equations to adequately relate the recorded performance characteristics. Consequently, every attempt is made to approximate performance characteristics by linear algebraic or linear differential equations.

In many instances the terminal equation can be established from manufacturer's specifications or derived from geometry (length, diameter, etc.) of the component. It is not our purpose, however, to discuss such procedures here.

The mechanical lever in Fig. 2.8a is considered as a three terminal component. Only vertical displacements and forces are to be considered with positive displacement $\delta(t)$ indicated by the arrow. The terminal graph in Fig. 2.8b indicates that two sets of variables are to be related by terminal equations. $\delta_1(t)$ is the instantaneous vertical displacement of point A with respect to the fulcrum point G and $f_1(t)$ the corresponding vertical force applied between the fixed reference G and point A. The variables $\delta_2(t)$ and $f_2(t)$ are similarly defined except they are measured relative to points B and G. When the lever characteristics are approximated by linear differential equations, the terminal equations may take the form

$$\begin{aligned} \delta_1(t) &= F_{11}(\tfrac{d}{dt}) \cdot f_1(t) + F_{12}(\tfrac{d}{dt}) \cdot \delta_2(t) \\ f_2(t) &= F_{21}(\tfrac{d}{dt}) \cdot f_1(t) + F_{22}(\tfrac{d}{dt}) \cdot \delta_2(t) \end{aligned} \qquad (2.1)$$

where $F_{ij}(\frac{d}{dt})$, (i, j = 1, 2) represent differential operators applied to the indicated variables. A more compact matrix notation for (2.1) is

$$\begin{bmatrix} \delta_1(t) \\ f_2(t) \end{bmatrix} = \begin{bmatrix} F_{11}(\frac{d}{dt}) & F_{12}(\frac{d}{dt}) \\ F_{21}(\frac{d}{dt}) & F_{22}(\frac{d}{dt}) \end{bmatrix} \begin{bmatrix} f_1(t) \\ \delta_2(t) \end{bmatrix} \qquad (2.2)$$

The differential operators appearing in the coefficient matrix of (2.2) can be established by considering the conceptual measurement apparatus of Fig. 2.9. The operators $F_{11}(\frac{d}{dt})$ and $F_{21}(\frac{d}{dt})$ are obtained by setting $\delta_2(t) = 0$ in (2.1) or (2.2) and allowing $f_1(t)$ to vary. The equivalent physical operation is shown in Fig. 2.9(a) where the lever is held in a horizontal position with end B fixed by a force meter measuring $f_2(t)$. The lever is excited by a force driver as indicated through the force meter measuring $f_1(t)$. Under these conditions $\delta_1(t)$ will also be zero, provided that the lever can be considered rigid. It is evident that $f_1(t)$ and $f_2(t)$ are related by

$$f_2(t) = n\, f_1(t)$$

where $n = \frac{\ell_1}{\ell_2}$. Thus we find that

$$F_{11}\left(\frac{d}{dt}\right) = 0 \text{ and } F_{21}\left(\frac{d}{dt}\right) = n$$

To evaluate $F_{12}\left(\frac{d}{dt}\right)$ and $F_{22}\left(\frac{d}{dt}\right)$, set $f_1(t) = 0$ (allow A to swing freely) and drive end B with a displacement driver as shown in Fig. 2.9b. For a rigid beam with rotational mass M and negligible friction we have

$$\delta_1(t) = -n\delta_2(t)$$

and

$$f_2(t) = M\frac{d^2}{dt^2}\delta_2(t)$$

Thus, the terminal equations modeling the performance characteristics of the lever are

$$\delta_1(t) = -n\delta_2(t)$$

$$f_2(t) = nf_1(t) + M\frac{d^2}{dt^2}\delta_2(t)$$

or in matrix form

$$\begin{bmatrix} \delta_1(t) \\ f_2(t) \end{bmatrix} = \begin{bmatrix} 0 & -n \\ n & M\frac{d^2}{dt^2} \end{bmatrix} \begin{bmatrix} f_1(t) \\ \delta_2(t) \end{bmatrix} \tag{2.3}$$

When the mass of the lever is negligible, M = 0 in (2.3) and the component is described as an "ideal" lever. The terminal equations in (2.3) with M = 0 are representative of a class of so-called "perfect coupler" components. Other examples of perfect couplers are electrical transformers with negligible losses and gear trains with negligible bearing friction and rotational mass.

Since the hydraulic control valve in Fig. 2.1 translates the mechanical variables of force and displacement measured between B and G to hydraulic variables of pressure and flow rate between C and D, it is referred to as a two-port transducer. Figure 2.10a shows the control valve together with a hydraulic pump where the two devices can be considered as a single two-port component. The terminal graph in Fig. 2.10b identifies each of the ports and the associated pairs of complementary variables. The terminal graph is separated into two parts since two different types of variables are identified.

The terminal equations for the valve depend upon those characteristics of the valve which must be taken into consideration. A reasonable linear model results when the variables are interrelated by equations of the form

$$\begin{bmatrix} f_1(t) \\ \dot{g}_2(t) \end{bmatrix} = \begin{bmatrix} (B_1\frac{d}{dt} + M_1\frac{d^2}{dt^2}) & 0 \\ K_{21} & K_{22} \end{bmatrix} \begin{bmatrix} \delta_1(t) \\ P_2(t) \end{bmatrix} \tag{2.4}$$

These equations include the mass and friction characteristics of the spool but assume that the pressure difference at points C and D does not affect the motion of the spool. The constants K_{21} and K_{22} depend upon the valve dimensions and the pressure applied to the valve by the hydraulic pump; the pump being of constant pressure type.

In a similar manner, it can be shown that the linear differential equations approximating the performance of the hydraulic ram in Figs. 2.1 and 2.11a are of the form

$$\begin{bmatrix} f_6(t) \\ \dot{g}_5(t) \end{bmatrix} = \begin{bmatrix} (B_6 \frac{d}{dt} + M_6 \frac{d^2}{dt^2}) & K_{56} \\ (-K_{65} \frac{d}{dt}) & 0 \end{bmatrix} \begin{bmatrix} \delta_6(t) \\ p_5(t) \end{bmatrix} \tag{2.5}$$

where the variables in the model, of course, are identified by the terminal graph of two parts in Fig. 2.11b.

In general, a mathematical model of an n-terminal electrical, mechanical or hydraulic component consists of: 1) a terminal graph of p parts having n vertices corresponding to the n terminals of the component and (n-p) oriented line segments connecting the n vertices, and 2) a set of (n-p) equations relating the 2(n-p) variables identified by the terminal graph.

3. Basic Graph Theory and Formulation of the State Model

The process of deriving or formulating a mathematical model of a system of interconnected components can best be described as a uniting of two sets of equations; this union producing a third set of equations called the system model. The two sets of equations from which the final model is evolved are derived from two separate properties of the system. One set, called the component equations, is simply a list of all terminal equations of the components in the system. They represent the characteristics of the individual system components and do not depend on the manner in which the components are connected together. The interconnection of the components (the system topology) is described by the second set of equations called interconnection or constraint equations. This set is always algebraic in form and relates the through and across variables at interfaces where the components are connected together. It is generally difficult to write the interconnection equations directly from a picture or schematic of the system. An examination of Fig. 2.1, for example, does not immediately reveal how many independent relationships are available and which variables should be related. This information is more efficiently obtained from a collection of oriented line segments called the system graph.

The system graph is formed by simply connecting together the terminal graphs from the mathematical models of the components in a one-to-one correspondence to the actual physical connections of the components in the system. For example, the system graph for the hydraulic subassembly shown in Figs. 2.1 and 2.3 is obtained by uniting the terminal graph of the control valve and pump assembly in Fig. 2.10b with a terminal graph for each of the hydraulic lines given in Fig. 2.5b and the terminal graph of the hydraulic ram given in Fig. 2.11b. The result is the collection of line segments shown in Fig. 3.1.

The oriented line segments in the system graph are called edges. The end points of the edges are called vertices and a set of edges forming a closed path is identified as a circuit. In this particular example, the graph contains 8 vertices, 6 edges, and only one circuit. In this respect it is a very simple system graph.

A simple graph having more than one circuit is shown in Fig. 3.2. In this case, three distinct circuits are formed by the sets of edges (3,4,5), (5,6,7) and (3,4,6,7). The set of edges (1,2) forms a path between vertices a and b while three distinct paths exist between vertices f and e; (5), (3,4) and (7.6). The subgraph (1,2,3,4) is not a path. A graph is said to be connected if there exists at least one path between every pair of vertices in the graph. All other graphs are separated graphs. The graph of Fig. 3.2 is a separated graph. It contains two connected subgraphs (1,2) and (3,4... 7). The subgraph (4,5,7) is connected while the subgraph (4,7) is not.

The two fundamental postulates of system theory state that 1) the summation of across variables associated with the edges of a connected graph vanish around any circuit and 2) the summation of the through variables vanish at the vertices. Thus, in the connected subgraph (3,4,...7) of Fig. 3.2 we have for the three circuits

$$\begin{aligned} &x_3 + x_7 - x_6 - x_4 = 0 \\ &x_3 + x_5 - x_4 = 0 \\ &x_5 + x_6 - x_7 = 0 \end{aligned} \tag{3.1}$$

and for the four vertices

$$\begin{aligned} &y_3 + y_4 = 0 \\ &y_4 + y_5 - y_6 = 0 \\ &y_7 + y_6 = 0 \\ &y_3 - y_5 - y_7 = 0 \end{aligned} \tag{3.2}$$

The reader can easily verify that one of the three circuit equations in (3.1) can be obtained as a linear combination of the other two and one of the four vertex equations in (3.2) can be obtained as a linear combination of the remaining three. The equations are, therefore, not independent.

A basis for selecting independent circuit and vertex equations is provided in the concept of a tree. A <u>tree</u> of a connected graph is defined as a connected subset of edges that contains all vertices of the graph but contains no circuits. Each connected subgraph of a graph contains at least one tree. For example, the connected subgraph (1,2) in Fig. 3.2 contains only one tree and it is the subgraph itself. On the other hand, the connected subgraph (3,4,5,6,7) contains several trees; (3,4,7), (3,7,6), (4,6,7), (3,4,6), (4,5,7), (3,5,7), (4,5,6) and (3,5,6). Subgraph (3,4,5,7) is not a tree because it contains a circuit. Likewise, subgraphs (3,6) and (3,5) cannot be trees because the former is not connected and the latter does not contain all of the vertices of the subgraph.

A collection of trees is appropriately called a <u>forest</u>, and a forest of a graph contains exactly <u>one</u> tree from each connected subgraph. The subgraph (1,2,3,4,7) comprises one forest of the graph in Fig. 3.2 and it is possible to define seven other distinct forests in the graph. After leading the reader this far into the woods, there should be no argument if the edges of a tree are called <u>branches</u>.

For every tree defined in a connected graph, there is a corresponding subgraph called a <u>co-tree</u> which contains those edges of the graph not contained in the tree. The edges of a co-tree are called <u>chords</u>. As a usual order of events, a tree of the graph is selected first and the corresponding co-tree is thus defined. In the connected subgraph (3,4,5,6,7) of Fig. 3.2 for example, the tree (4,5,6) defines a co-tree with chords 3 and 7. If the tree of a graph is the graph itself, as in the subgraph (1,2), then the co-tree is an empty set.

If a connected graph has <u>e</u> edges and <u>v</u> vertices, then as a result of the definitions above, <u>any</u> tree of the graph always has <u>v-1 branches</u> and <u>e-v+1 chords</u>. The connected subgraph (3,4,5,6,7) has 5 edges and 4 vertices. Thus every tree of this graph must have (v-1) = 3 branches and (e-v+1) = 2 chords.

One important property of a tree is that the addition of a chord between any two vertices of the tree establishes a circuit. This is evident when the edges of a tree are identified by heavy lines as in Fig. 3.3a. The tree indicated defines a co-tree with chords 5 and 7. Each of these chords, when added to the tree defines a <u>unique</u> circuit composed of the defining chord and a subset of the branches of the tree.

The circuits defined by the chords of a tree are called <u>fundamental circuits</u> of a graph. Since a defining chord appears in <u>only one</u> fundamental circuit they are unique and form a basis for writing <u>independent</u> circuit equations from the graph. For example, the two fundamental circuit equations for the graph of Fig. 3.3a are

$$\begin{aligned} x_5 - x_4 + x_3 &= 0 \\ x_7 - x_6 - x_4 + x_3 &= 0 \end{aligned} \tag{3.3}$$

and we see that the two equations are independent since each equation includes an across variable not contained in the other. The two circuit equations in (3.3) can also be written as

$$\begin{aligned} x_5 &= x_4 - x_3 \\ x_7 &= x_6 + x_4 - x_3 \end{aligned} \tag{3.4}$$

It can be shown that, as in the above example, the circuit equations for the system graph can always be written so as to show the <u>across variables of the chords as an explicit function of the across variables of the branches</u>. Such circuit equations are called the <u>fundamental circuit equations</u> and are written after a tree has been selected in the system graph. The number of such independent circuit equations, of course, is equal to the number of chords in the graph, i.e., (e-v+1) if the graph is connected and has <u>e</u> edges and <u>v</u> vertices.

A similar criterion for obtaining independent relationships among the through variables follows from another interesting property of the tree. Let S be a set of edges of a connected system graph which has the following properties: (1) removal of the set S leaves the graph in <u>exactly two parts</u> (an isolated vertex qualifies as a part) and (2) the set S contains <u>exactly one branch</u> of a tree plus the <u>minimum</u> number of chords necessary to satisfy property (1). Such a set of edges is called a <u>fundamental cut-set</u> or simply a cut-set. Referring to the graph of Fig. 3.3b, the three cut-sets defined by the branches of the tree are indicated by the dashed lines I, II, and III and consist, respectively, of edges (3,5,7), (4,5,7) and (6,7). It is evident that each cut-set is <u>unique</u> in that each branch of the tree appears in <u>exactly one</u> cut-set.

An almost obvious generalization of the vertex equations is that the summation of the through variables associated with a cut-set of edges vanishes identically. The three cut-set equations for the graph of Fig. 3.3b, for example, are

$$\begin{aligned} y_3 - y_5 - y_7 &= 0 \\ y_4 + y_5 + y_7 &= 0 \\ y_6 + y_7 &= 0 \end{aligned} \tag{3.5}$$

These equations are clearly independent since each one contains a variable not contained in either of the others. This implies also that the three cut-set equations in (3.5) can also be written as

$$\begin{aligned} y_3 &= y_5 + y_7 \\ y_4 &= - y_5 - y_7 \\ y_6 &= - y_7 \end{aligned} \tag{3.6}$$

Note that these equations show the <u>branch through variables as an explicit function of the chord through variables</u>. Such a system of equations is called the <u>fundamental cut-set</u> equations. It is not difficult to see that since each cut-set contains exactly one branch of the tree, the cut-set equations for a graph can always be written in this form and that the number of such equations is exactly equal to the number of branches in the graph, i.e., (v-1) if the graph is connected and contains <u>v</u> vertices.

It might appear that the circuit equations given in (3.4) and the cut-set equations given in (3.6) are entirely unrelated by virtue of the separate methods used to derive them. To the contrary, they are very much related. But the relationship between them is only evident when (3.4) and (3.6) are written in the matrix forms

$$\begin{bmatrix} x_5 \\ x_7 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_3 \\ x_4 \\ x_6 \end{bmatrix} \tag{3.7}$$

and

$$\begin{bmatrix} y_3 \\ y_4 \\ y_6 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & -1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} y_5 \\ y_7 \end{bmatrix} \tag{3.8}$$

The arrays of ± 1 and 0 coefficients to the right of the equality signs are called the fundamental circuit and cut-set matrices, respectively. Note that one coefficient matrix is the negative transpose of the other, i.e., one matrix has two rows and three columns while the other has three rows and two columns and the entries in the <u>rows</u> of one matrix are identical to the negative of the entries in the <u>columns</u> of the other. The coefficient matrices of the fundamental circuit and cut-set equations of any system graph are related in this manner. Thus, if the vectors X_b and X_c represent the across variables associated, respectively, with the branches and chords of a system graph, and if Y_b and Y_c are the corresponding vectors of through variables, then (3.7) and (3.8) can be regarded as specific examples of the general system of fundamental circuit and cut-set equations.

$$X_c = BX_b \tag{3.9}$$

and

$$Y_b = AY_c \tag{3.10}$$

where A and B are rectangular matrices and $A = -B^T$, i.e. A is the negative transpose of B.

Among other things, (3.9) and (3.10) imply conservation of energy in any closed system. To show this, note that from (3.9) and (3.10) it follows that

$$X_e = \begin{bmatrix} X_b \\ \hline X_c \end{bmatrix} = \begin{bmatrix} U \\ \hline B \end{bmatrix} X_b \tag{3.11}$$

$$Y_e = \begin{bmatrix} Y_b \\ \hdashline Y_c \end{bmatrix} = \begin{bmatrix} A \\ \hdashline U \end{bmatrix} Y_c = \begin{bmatrix} -B^T \\ \hdashline U \end{bmatrix} Y_c \tag{3.12}$$

where X_e and Y_e are vectors containing all across and through variables, respectively, and where U represents the unit or identity matrix. The total power (or energy) in the system is given by the scalar product, or dot product of the vectors X_e and Y_e written from 3.11 and 3.12 as

$$X_e^T Y_e = X_b^T [U \;\vdots\; B^T] \begin{bmatrix} -B^T \\ \hdashline U \end{bmatrix} Y_c = X_b^T(-B^T + B^T)Y_c \tag{3.13}$$

$$= 0$$

For the circuit and cut-set equations in (3.7) and (3.8) the expression symbolized in (3.13) is specifically

$$[x_5\ x_7 \;\vdots\; x_3\ x_4\ x_6] \begin{bmatrix} y_5 \\ y_7 \\ \hdashline y_3 \\ y_4 \\ y_6 \end{bmatrix} = [x_3\ x_4\ x_6] \left[\begin{array}{ccc:cc} 1 & 0 & 0 & -1 & -1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{array}\right] \begin{bmatrix} 1 & 1 \\ -1 & -1 \\ 0 & -1 \\ \hdashline 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_5 \\ y_7 \end{bmatrix}$$

$$= [x_3\ x_4\ x_6] \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_5 \\ y_7 \end{bmatrix}$$

$$= 0$$

In summary, if a connected graph contains v vertices and e edges, there are always e interconnection equations; (v-1) cut-set equations expressing the branch through variables and (e-v+1) circuit equations expressing the chord across variables in terms of the branch across variables. Thus, the interconnection equations of a system provide e independent equations in 2e unknowns. These e equations combine with the e component equations, (in the same variables) characterizing the performance characteristics of the components, to provide 2e equations in 2e variables. This set of equations constitutes the basis for all mathematical models of the performance characteristics of the system.

The final explicit form of the mathematical model of a system varies greatly, depending on the form of the equations used to model the characteristics of the components, the type of information to be gained from the analysis and the "machinery" available for solving the model. The so-called state model is easily derived for most systems and is exceptionally well suited for machine solution by analog or digital computer.

Considering the hydraulic lift system of Fig. 2.1, the state model of the entire system can be formulated in one operation but it is simpler to consider the system as a collection of three subassemblies as in Fig. 2.2. In this way, the complete system state model is derived from the three individual models; again considering the system as the sum of its parts.

Consider first the hydraulic subassembly shown in Fig. 2.3. We desire a mathematical model of the input-output performance characteristics of the assembly as a two-port component. The models of the various components in this subassembly and the resulting system graph have already been discussed in detail with the system graph shown explicitly in Fig. 3.1. This graph is repeated in Fig. 3.4a with a tree indicated by heavy lines. The required mathematical model is obtained by using the fundamental circuit and cut-set equations of the system graph to derive a set of equations showing $f_1(t)$ and $\delta_6(t)$ as a function of $\delta_1(t)$ and $f_6(t)$ only. Note that $\delta_1(t)$ and $\delta_6(t)$ are not in general measured with respect to the <u>same</u> fixed reference G. Since the <u>relative</u> time variation between $\delta_1(t)$ and $\delta_6(t)$ is of primary interest, the derived mathematical model will not account for the constant distance between the two fixed references. In the interest of simplifying the algebra in this example development, the mass and friction of the control valve are neglected and the hydraulic lines are modeled as pure resistors. With these simplifications, the component terminal equations (2.4) and (2.5) together with the terminal equations for the hydraulic lines can be listed as follows:

$$\frac{d}{dt}\begin{bmatrix}\dot{\delta}_6\\ \delta_6\end{bmatrix} = \begin{bmatrix}(-B_6/M_6)\dot{\delta}_6 - (K_{56}/M_6)p_5 + (1/M_6)f_6\\ \dot{\delta}_6\end{bmatrix} \tag{3.14}$$

$$\begin{bmatrix}p_2\\ p_3\\ p_4\\ \dot{g}_5\end{bmatrix} = \begin{bmatrix}(-K_{21}/K_{22})\ \delta_1 + (1/K_{22})\ \dot{g}_2\\ R_3\ \dot{g}_3\\ R_4\ \dot{g}_4\\ -K_{65}\ \dot{\delta}_6\end{bmatrix} \tag{3.15}$$

Note that the component equations have been divided into two groups; the differential equations and the algebraic equations. Note further that the differential equations have been written in a form explicit in the first derivative of the variable. In addition, the algebraic equations are explicit in the across variables if, in the system graph of Fig. 3.4a, the corresponding edges are branches of the tree and explicit in the through variable if the edge is a chord.

The second step in developing the required model of the subassembly is to use the fundamental circuit and cut-set equations of the system graph in Fig. 3.4a to eliminate all branch through variables and all chord across variables from the component equations in (3.14) and (3.15). In this particular example, the interconnection equations are simply

$$\begin{gathered} p_5 = -p_3 + p_2 + p_4 \\ \dot{g}_3 = \dot{g}_5,\ \dot{g}_4 = -\dot{g}_5,\ \dot{g}_2 = -\dot{g}_5 \end{gathered} \tag{3.16}$$

substituting (3.16) into (3.14) and (3.15) gives

$$\frac{d}{dt}\begin{bmatrix} \dot{\delta}_6 \\ \delta_6 \end{bmatrix} = \begin{bmatrix} (-B_6/M_6)\dot{\delta}_6 - (K_{56}/M_6)(p_2 - p_3 + p_4) + (1/M_6)f_6 \\ \dot{\delta}_6 \end{bmatrix} \tag{3.17}$$

$$\begin{bmatrix} 1 & 0 & 0 & 1/K_{22} \\ 0 & 1 & 0 & -R_3 \\ 0 & 0 & 1 & R_4 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} p_2 \\ p_3 \\ p_4 \\ \dot{g}_5 \end{bmatrix} = \begin{bmatrix} (-K_{21}/K_{22})\delta_1 \\ 0 \\ 0 \\ -K_{65}\dot{\delta}_6 \end{bmatrix} \tag{3.18}$$

The third, and last, step in developing the required model of the subassembly is to solve the algebraic system of equations in (3.18) for p_2, p_3 and p_4 and substitute the result into the differential equations in (3.17). The solution in this particular case is very simple and the resulting model of the system is

$$\frac{d}{dt}\begin{bmatrix} \dot{\delta}_6 \\ \delta_6 \end{bmatrix} = \begin{bmatrix} -\frac{B_6}{M_6} - \frac{K_{56}K_{65}}{M_6}\left(\frac{1}{K_{22}} + R_3 + R_4\right)\dot{\delta}_6 + \frac{K_{56}K_{21}}{M_6 K_{22}}\delta_1 + \frac{1}{M_6} f_6 \\ \dot{\delta}_6 \end{bmatrix} \tag{3.19}$$

or

$$\frac{d}{dt}\begin{bmatrix} \dot{\delta}_6 \\ \delta_6 \end{bmatrix} = \begin{bmatrix} A_h & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} \dot{\delta}_6 \\ \delta_6 \end{bmatrix} + \begin{bmatrix} \frac{(K_{56}K_{21})}{M_6 K_{22}} & \frac{1}{M_6} \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \delta_1 \\ f_6 \end{bmatrix}$$

where

$$A_h = -(B_6/M_6) - (K_{56}K_{65}/M_6)(1/K_{22} + R_3 + R_4) \tag{3.20}$$

The final model in (3.19) has the form of two simultaneous differential equations in the unknowns $\dot{\delta}_6$ and δ_6. The equations are first order nonhomogenous and are linear under the assumption that the coefficient matrices have constant entries. For any specified time variation of the terminal variables (drivers) δ_1 and f_6, the solution to (3.19) yields $\dot{\delta}_6$ and δ_6 as functions of time and the specified drivers. The "state" of the system is dependent upon $\dot{\delta}_6$ and δ_6 since the remaining variables in the subassembly are determined by the appropriate relationships in (3.18) and/or the interconnection equations. For this reason, the variables $\dot{\delta}_6$ and δ_6 are called <u>state variables</u> and the differential equations the <u>state equations</u>. These equations together with the trivial terminal equation $f_1 = 0$ comprise the <u>state model</u> of the subassembly. Combining the terminal graph of Fig. 3.4b and the state model, we have a <u>complete mathematical model of the component</u>; thus allowing the hydraulic subassembly to be regarded as one two-port component of the larger system.

In a similar manner, it can be shown that the state model of the lever and mass load in Fig. 3.5a, considered as a three-terminal component, is

$$\frac{d}{dt}\begin{bmatrix}\dot{\delta}_8\\ \delta_8\end{bmatrix} = \begin{bmatrix}0 & (-1/M_L K_s)\\ 1 & 0\end{bmatrix}\begin{bmatrix}\dot{\delta}_8\\ \delta_8\end{bmatrix} + \begin{bmatrix}(n_e/M_L K_s) & (1/M_L)\\ 0 & 0\end{bmatrix}\begin{bmatrix}\delta_7\\ f_8\end{bmatrix}$$

$$f_7 = (-n_e/K_s)\delta_8 + (n_e^2/K_s)\delta_7 \qquad (3.21)$$

where the variables in the state and terminal equations refer to the graph of Fig. 3.5b. Also, $n_e = (\ell_2/\ell_1)$ represents the ratio of lever arms, K_s represents the spring constant of the lever and M_L represents the load mass. The mass of the lever has been neglected.

Assuming that the control lever AB in Fig. 2.1 is "ideal" its terminal equations and terminal graph are taken, respectively from (2.3) with M = 0 and Fig. 2.8b. A state model of the complete hydraulic lift system is obtained by first combining the terminal graph of the control lever with the terminal graphs in Figs. 3.4b and 3.5b, consistent with the union of the components, to establish the system graph of Fig. 3.6. The unoriented line is simply an expanded vertex, edge 11 identifies the independent input variable δ_{11}, edges 9 and 10 identify, respectively, the input and output variables of the control lever, edges 1 and 6 identify the input and output variables of the hydraulic subassembly and 7 and 8 identify the input and output of the mass-lever subassembly. The respective component graphs in this case are said to be connected in cascade.

A list of the terminal equations for the components in the system is

$$\frac{d}{dt}\begin{bmatrix}\dot{\delta}_6\\ \delta_6\\ \dot{\delta}_8\\ \delta_8\end{bmatrix} = \begin{bmatrix}A_h \dot{\delta}_6 + \dfrac{K_{56}K_{21}}{M_6 K_{22}}\delta_1 + \dfrac{1}{M_6} f_6\\ \dot{\delta}_6\\ -\dfrac{1}{M_L K_s}\delta_8 + \dfrac{n_e}{M_L K_s}\delta_7 + \dfrac{1}{M_L} f_8\\ \dot{\delta}_8\end{bmatrix} \qquad (3.22)$$

and

$$\begin{bmatrix}\delta_{10}\\ f_7\\ f_9\end{bmatrix} = \begin{bmatrix}-n_c\delta_9\\ -\dfrac{n_e}{K_s}\delta_8 + \dfrac{n_e^2}{K_s}\delta_7\\ -n_c f_{10}\end{bmatrix} \qquad (3.23)$$

where n_c represents the lever arm ratio of the control lever. Note that the tree in the system graph has been selected so that the algebraic equations can be written explicit in branch displacements and chord forces or their time derivatives.

As in the development of the state model of the hydraulic subassembly, the state model of the entire hydraulic lift system is obtained by: 1) eliminating all branch forces and chord displacements from (3.22) and (3.23) by applying the fundamental circuit and cut-set equations for the tree indicated in the system graph, 2) solve the resulting algebraic system of equations in (3.23) for δ_{10} and f_7 and 3) substitute this solution into the differential equations. The result is the set of first-order differential equations characterizing the performance of the system.

$$\frac{d}{dt}\begin{bmatrix} \dot{\delta}_6 \\ \delta_6 \\ \dot{\delta}_8 \\ \delta_8 \end{bmatrix} = \begin{bmatrix} A_h & \frac{-n_e^2}{M_6 K_s} & 0 & \frac{n_e}{M_6 K_s} \\ 1 & 0 & 0 & 0 \\ 0 & \frac{n_e}{M_L K_s} & 0 & \frac{-1}{M_L K_s} \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \dot{\delta}_6 \\ \delta_6 \\ \dot{\delta}_8 \\ \delta_8 \end{bmatrix} + \begin{bmatrix} \frac{-n_c K_{21} K_{56}}{M_6 K_{22}} & 0 \\ 0 & 1 \\ 0 & \frac{1}{M_L} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \delta_{11} \\ f_8 \end{bmatrix} \tag{3.24}$$

$$f_{11} = 0$$

The differential equations are the state equations of the system since all variables in the system are obtainable as a linear combination of the variables appearing in the state vector. In this particular case, the two input variables are the control lever position δ_{11} and the externally applied load force f_8. If no additional load force is applied then $f_8 \equiv 0$. The output variables of primary interest is the load position δ_8 and the force f_{11} required to move the control lever. The latter force is identically zero under the assumptions used to model the control valve and δ_8 appears explicitly as one of the variables in the state vector.

4. Solution of the State Model

All performance characteristics of the hydraulic lift system are implied in the solution of the equations of state given in (3.24). Linear state models of this type can be solved analytically to obtain the output variables. In addition, they may be solved numerically by either analog or digital computers to obtain a plot of the output for any particular specified time variations in the input variables. Thus, the solution for a state variable $z_1(t)$ might typically be given in any one of the three forms shown in Fig. 4.1.

It is evident that the analytic solution is the most desirable since the other two solutions can be written directly from it. On the other hand, the exact analytic solution can never be found from a numerical solution although under certain conditions it can be approximated. As might be expected the analytic solution, as the most desirable type, is frequently the most difficult to obtain. In fact, an analytic solution may not be obtainable in a usable form for some state models, particularly nonlinear ones.

The use of analog computers in the solution of state models is considered by Hammond. The emphasis here is, therefore, on analytical procedures. These procedures fall into two basic classes; Laplace transform methods and functions of matrices. Both of these procedures are considered briefly together with a short discussion of a digital computer method based on functions of matrices.

A widely-used analytical method for solving ordinary linear differential equations is based on the Laplace transform. Many useful design concepts and procedures employ the so-called complex frequency models. These models can be obtained by directly applying the Laplace transform method to the state equations. The reader unfamiliar with the Laplace transform will find it instructive to supplement the discussion of the subject given here with reading in almost any modern book on feedback systems. Our purpose here is to show the application of the Laplace transform to the solution of the state model and derivation of the complex frequency model from the state model.

If F(t) is a vector of time functions

$$F(t) = \begin{bmatrix} f_1(t) \\ f_2(t) \\ \cdot \\ \cdot \\ f_n(t) \end{bmatrix} \tag{4.1}$$

then the Laplace transform of F(t), written L[F(t)], is defined to be the vector

$$L[F(t)] = F(s) = \begin{bmatrix} L[f_1(t)] \\ L[f_2(t)] \\ \cdot \\ \cdot \\ L[f_n(t)] \end{bmatrix} = \begin{bmatrix} f_1(s) \\ f_2(s) \\ \cdot \\ \cdot \\ f_n(s) \end{bmatrix} \tag{4.2}$$

where

$$L[f_i(t)] = f_i(s) = \int_0^{\infty} f_i(t)e^{-st}dt, \quad i = 1, 2, \ldots, n \tag{4.3}$$

and $\underline{s}$ is a complex variable $s = \sigma + j\omega$ sometime called the <u>complex frequency</u>. It is assumed, of course that $f_i(t)$ in the definition of (4.3) is such that the indicated integral exists. A short table of time functions and their corresponding Laplace transforms is given in Table 4.2.

It is easy to show through integration by parts that if the Laplace transform of a vector Z(t) is Z(s) then the Laplace transform of $\frac{d}{dt} Z(t)$ is given by

$$L\left[\frac{d}{dt} Z(t)\right] = sZ(s) - Z(0+) \tag{4.4}$$

where $Z(0+) = Z_o = \lim_{t \to 0} Z(t)$ with t assuming only positive values.

Thus, Z_o represents the initial value of the vector.

Consider now the state model

$$\frac{d}{dt} Z(t) = AZ(t) + BF(t) \tag{4.5}$$

$$E(t) = CZ(t) + DF(t) \tag{4.6}$$

Where $F(t)$ and $E(t)$ are vectors representing the input and output variables respectively, $Z(t)$ is an n-dimensional state vector and A, B, C and D are constant matrices of appropriate dimensions. Note that the state model developed in the previous section is a degenerate from of the general model in (4.5) and (4.6).

Taking the Laplace transform of (4.5) and (4.6) and applying (4.4) gives the so-called complex frequency or <u>s-domain</u> <u>model</u>.

$$sZ(s) - Z_o = AZ(s) + BF(s) \tag{4.7}$$

$$E(s) = CZ(s) + DF(s) \tag{4.8}$$

Note that the transformation has converted the differential equations in (4.5) to an algebraic system of equations. Consequently, it is now possible to solve (4.7) for $Z(s)$. Collecting the coefficients of $Z(s)$ in (4.7) yields

$$(sU - A)\, Z(s) = Z_o + BF(s) \tag{4.9}$$

where U is an nxn unit matrix. The coefficient matrix $(sU - A)$ is called the <u>characteristic</u> matrix and its inverse can be written as

$$(sU - A)^{-1} = \frac{\text{Adj}\,(sU - A)}{d(s)} = \frac{P(s)}{d(s)} \tag{4.10}$$

$P(s)$ is the <u>adjoint</u> <u>matrix</u> of $(sU - A)$ and has as its entries polynomials in <u>s</u> of degree $(n-1)$ or less. The polynomial $d(s)$ is called the <u>characteristic</u> <u>polynomial</u> of A and represents the determinant of $(sU - A)$. Solving (4.9) for $Z(s)$ by multiplying both sides by $(sU - A)^{-1}$ defined in (4.10) we obtain

$$Z(s) = \frac{1}{d(s)} P(s)\, Z_o + \frac{1}{d(s)} P(s)\, BF(s) \tag{4.11}$$

For any given $F(t)$, $F(s)$ is a function of s and the vector $Z(s)$ as given by (4.11) represents the Laplace transform of some vector function of time. The only question is <u>what</u> function of time. If this can be determined then the solution to the differential equations in the state model has been determined. The process of going back from a known function of <u>s</u> to the corresponding function of time is called the inverse Laplace transform and is best achieved by expanding the ratios of polynomials on the right-hand side of (4.11) in partial fractions. But to realize this expansion, the zeros of $d(s)$ must be known. Let $d(s)$ be written in the factored form

$$d(s) = (s - s_1)(s - s_2) \cdots (s - s_n) \tag{4.12}$$

Assuming that the zeros s_1, s_2, -- s_n are distinct, $Z(s)$ can be expanded in partial fractions as

$$Z(s) = \frac{R_1}{(s - s_1)} + \frac{R_2}{(s - s_2)} + \ldots + \frac{R_n}{(s - s_n)} + \frac{F_1}{(s - s_1)} + \frac{F_2}{(s - s_2)} +$$

$$+ \ldots + \frac{F_n}{(s-s_n)} + \frac{F_a}{(s-s_a)} + \frac{F_b}{(s-s_b)} + \ldots \frac{F_k}{(s-s_k)} \tag{4.13}$$

where the constant vectors R_i and F_i are given by

where the constant vectors R_i and F_i are given by

$$R_i = \lim_{s \to s_i} \frac{(s-s_i)}{d(s)} P(s)Z_o, (i = 1, 2, \ldots, n)$$

$$F_i = \lim_{s \to s_i} \frac{(s-s_i)}{d(s)} P(s)BF(s), (i = 1, 2, \ldots n, a, b..k)$$

and s_a, s_b, ..., s_k are distinct poles contributed by the vector $F(s)$. By definition, a vector $F(s)$ has a pole at $s = s_p$ if for at least one entry $f_i(s)$ in $F(s)$ we have

$$\lim_{s \to s_p} f_i(s) = \infty \tag{4.14}$$

Each factor in the partial fraction expansion in (4.13) corresponds to the Laplace transform pair:

$$F(t) = K e^{\alpha t} \Longrightarrow F(s) = \frac{K}{(s-\alpha)} \tag{4.15}$$

where K is a constant matrix (or vector) and α is a constant scalar. Applying this equivalence to (4.13), it is evident that the state vector $Z(t)$ which has $Z(s)$ as its Laplace transform is

$$Z(t) = G_1 e^{s_1 t} + G_2 e^{s_2 t} + \ldots + G_n e^{s_n t} + F_a e^{s_a t} + \ldots + F_k e^{s_k t} \tag{4.16}$$

where the vector $G_i = R_i + F_i$, $i = 1,2,\ldots,n$.

Thus, (4.16) represents one general form of the solution to the state equations of (4.5). The purely exponential character normally occurs only when the zeros of $d(s)$ are distinct. When repeated zeros occur, the solution may have terms of the form $G_i t^m e^{s_i t}$, $m = 1,2,\ldots n$. In general, the zeros of $d(s)$ can be real or complex and when complex they occur as conjugate pairs. A stable solution $Z(t)$ is one having bounded (finite) state variables as time increases without bound (assuming the input variables in $F(t)$ are bounded). An examination of (4.16) reveals that <u>$Z(t)$ is a stable solution if all zeros of $d(s)$, $s_1, s_2, \ldots s_n$ have negative real parts</u>. This follows immediately from the fact that $e^{s_j t} \to 0$ as $t \to \infty$ when the real part of $s_j < 0$, $j = 1,2,\ldots n$. Further discussion of stability and a detailed example of a solution by Laplace transform methods is given in Section 5.

The complex-frequency or s-domain model of the system is, by definition, obtained by substituting the solution of (4.7) for $Z(s)$ into (4.8) <u>with $Z_o = 0$</u>. The result is

$$E(s) = [C(sU - A)^{-1}B + D]F(s) = \bar{G}(s)F(s) \tag{4.17}$$

where $\bar{G}(s) = [C(sU - A)^{-1}B + D]$ is a <u>matrix</u> whose entries are ratios of polynomials. Note that the initial conditions are taken as zero. This is always implied in a given s-domain model. Thus, the s-domain mathematical model of a typical two-port component takes the form

$$\begin{bmatrix} y_1(s) \\ x_2(s) \end{bmatrix} = \begin{bmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{bmatrix} \begin{bmatrix} x_1(s) \\ y_2(s) \end{bmatrix} \tag{4.18}$$

where $G_{ij}(s)$, $(i,j = 1,2)$, is a ratio of polynomials in $\underline{s}$. If port 1 is identified as the input port and 2 as the output port, then $G_{11}(s)$ is called the complex frequency input admittance, $G_{22}(s)$ is called the complex frequency output impedance, $G_{21}(s)$ the forward transfer function and $G_{12}(s)$ the reverse or return transfer function. The entire coefficient matrix in (4.18) is called the transfer matrix.

An important relationship exists between the forward transfer function of a system and the coefficient matrix A of the corresponding system state model. Using (4.10) in the definition of the matrix $\bar{G}(s)$ below (4.17), we can write

$$\bar{G}(s) = [\frac{CP(s)B}{d(s)} + D]$$

Since C, B and D are constant matrices, it follows that each ratio $G_{ij}(s)$ in the transfer matrix of (4.18) has part or all of the factors of d(s) in its denominator. Therefore, the forward transfer function $G_{21}(s)$ has the general form

$$G_{21}(s) = \frac{n(s)}{d(s)} = \frac{n(s)}{(s-s_1)(s-s_2)\ldots(s-s_n)}$$

where some of the factors of d(s) may not appear in the denominator of $G_{21}(s)$ by virtue of cancellation with n(s). The values s_1, s_2, .., s_n are poles of the transfer function $G_{21}(s)$ and it is obvious that these values satisfy the characteristic equation of the matrix A; $d(s) = 0$. Thus, under certain conditions, the stability of a system can be determined from the poles of its forward transfer function.

Example 4.1. The s-domain transfer characteristics of the hydraulic subassembly in Fig. 2.1 are obtained by applying the procedure described above to (3.19). In this case (4.18) has the particular form

$$\begin{bmatrix} f_1(s) \\ \delta_6(s) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ G_{61}(s) & G_{66}(s) \end{bmatrix} \begin{bmatrix} \delta_1(s) \\ f_6(s) \end{bmatrix} \tag{4.19}$$

$$\text{where } G_{61}(s) = \frac{K_{56}K_{21}}{M_6K_{22}(s^2 - A_h s)} \tag{4.20}$$

$$G_{66}(s) = \frac{1}{M_6(s^2 - A_h s)} \tag{4.21}$$

Note that the characteristic polynomial of the coefficient matrix in (3.19), $d(s) = \det(sU - A) = s^2 - A_h s$ appears in the denominator of $G_{61}(s)$ and $G_{66}(s)$. Also, in this particular case both the s-domain input admittance and return transfer function are identically zero. Such a component is said to be a unilateral component with zero input admittance--the output variable $f_6(s)$ has no effect on the input variable $f_1(s)$. In a special application where the load connected to the assembly is negligible, $f_6(s)$ is taken as zero and only the no-load transfer function $G_{61}(s)$ is of interest, and we write

$$\delta_6(s) = G_{61}(s)\,\delta_1(s) \tag{4.22}$$

Consequently, if two unilateral components are cascaded to form the system in Fig. 4.3a, then under the assumption of no loading, the system transfer function is given as the product of the component no-load transfer functions. The block diagram shown in Fig. 4.3b is used to indicate this very special but important result. The blocks and arrows imply that

$$x_4(s) = G_2(s)\,x_3(s) = G_2(s)\,G_1(s)\,x_1(s) \tag{4.23}$$

and therefore the system transfer function is, by definition

$$G(s) = \frac{x_4(s)}{x_1(s)} = G_2(s)\,G_1(s) \tag{4.24}$$

The transfer function and the associated block diagram are valuable tools in treating complex systems where no "loading" occurs between components, i.e., when the through variable measured at the interface between each pair of compoents is zero or can be assumed so. More discussion relating the transfer function to the concepts of feedback and stability appears in Section 5.

Another technique for obtaining the analytical solution to a linear state model is based on functions of matrices. A function of a matrix is a matrix istelf and can be evaluated in the same way as the scalar function $h = f(x)$. For example, let the matrix Y be the following polynomial function of the matrix X

$$Y = f(X) = U + 3X + 8X^2 - 15X^3 \tag{4.25}$$

Upon substituting a square matrix X into the right side of (4.25), the matrix Y can be evaluated.

To illustrate briefly the use of matrix functions in the solution of linear differential equations, consider the nohomogenous state equations

$$\frac{d}{dt} Z(t) = A\ Z(t) \tag{4.26}$$

with initial conditions $Z(0+) = Z_o$. Suppose the nxn matrix e^{At} is defined so that

$$\frac{d}{dt}(e^{At}) = A\,e^{At} \tag{4.27}$$

and

$$\lim_{t \to 0} e^{At} = U \tag{4.28}$$

where U is the nxn unit matrix. Under these conditions the solution of the state equations in (4.26) can be expressed as

$$Z(t) = e^{At}\ Z_o \tag{4.29}$$

By substituting (4.29) into (4.26), $Z(t)$ can be verified as a solution with initial conditions Z_o.

The matrix e^{At} is a function of the matrix A; this function being defined by the infinite matrix series

$$e^{At} = U + At + \frac{A^2t^2}{2!} + \frac{A^3t^3}{3!} + \dots \qquad (4.30)$$

Notice that the properties given in (4.27) and (4.28) follow directly from (4.30).

The series in (4.30) can be evaluated analytically, and in closed form in terms of constant matrices and exponentials, i.e.

$$e^{At} = C_1e^{\lambda_1 t} + C_2e^{\lambda_2 t} + \dots + C_ne^{\lambda_n t} \qquad (4.31)$$

where the constant matrices C_1, C_2, -- C_n are called <u>constituent</u> matrices and the constants λ_1, λ_2 --- λ_n are the <u>eigenvalues</u> of the matrix A; assumed here to be distinct. Substituting (4.31) into (4.29) yields the final analytic solution to the homogeneous state model as

$$Z(t) = R_1e^{\lambda_1 t} + R_2e^{\lambda_2 t} + \dots + R_ne^{\lambda_n t} \qquad (4.32)$$

where the vector $R_i = C_i Z_o$, $i = 1, 2, \dots, n$. While the procedure for obtaining the constituent matrices C_1, C_2 ... C_n is not difficult, a more detailed discussion of the method is beyond the scope of this paper. The interested reader should consult the attached list of references.

It should be noted that the method outlined above applies <u>only</u> to the <u>homogeneous</u> state model given by (4.26). However, the <u>non-homogeneous</u> model in (4.5) can be transformed into a homogeneous form when certain conditions on the input variables in $F(t)$ are satisfied. These conditions are not stringent; being roughly equivalent to the requirement that the Laplace transform of each variable exist.

In view of the fact that a later paper discusses the application of computers to the solution of differential equations, a detailed development is not given here. However, one particular method, based on the series representation of the matrix e^{At} in (4.30) deserves some comment. This method utilizes the digital computer and provides a discrete numerical solution of the homogeneous state model of (4.26). Although the method provides an approximate solution, the approximation error can be made as small as desired.

Let the interval of time over which a solution is desired be divided into small increments of length <u>h</u> as shown in Fig. 4.4. Let $z_i(t)$ be a typical state variable of the state vector $Z(t)$. With the time increment <u>h</u> specified, we approximate the matrix e^{Ah} by evaluating the first few terms of the series

$$e^{Ah} \doteq U + Ah + \frac{A^2h^2}{2!} + \frac{A^3h^3}{3!} + \dots \qquad (4.33)$$

The number of terms considered naturally depends upon the allowable solution error. From (4.29) we can write

$$Z(1) = Z(h) = e^{Ah}Z_o \qquad (4.34)$$

$Z(1)$ approximates the value of the state vector $Z(t)$ at time $t = h$. Letting $Z(1)$ be the new initial condition we can similarly write from (4.29) that

$$Z(2) = Z(2h) = e^{Ah} Z(1)$$
$$= (e^{Ah})^2 Z_o \tag{4.35}$$

Continuing in this manner, the value of the state vector at any discrete time $t = nh$ can be approximated by

$$Z(n) = e^{Ah} Z(n-1)$$
$$Z(n) = (e^{Ah})^n Z_o \tag{4.36}$$

The approximation can obviously be improved by reducing the size of the increment h.

The recursion formula given in (4.36) is easily programmed on the digital computer as is the original evaluation of e^{Ah} from (4.33). At each iteration, the matrix e^{Ah} is raised to the next highest power and multiplied into the given initial vector Z_o. This process yields a sequence of numbers for each of the state variables. Fig. 4.4 shows the resulting solution for a typical variable.

5. Feedback Systems

The concept of "feedback" is basic to the study of almost all control systems and is centered about a particular pattern of interconnection between components. Although deceptively simple to define, the effect of feedback on a given system may be complex. Suppose it is desired to accurately control the value of the output across variable $x_o(t)$ by varying the input $x_i(t)$. Theoretically, if the exact numerical correspondence between input and output values is known, the required input signal for a desired output can be calculated and the control problem is trivially solved on paper by an open-loop system. However, in practical systems, no exact relationship between the input and output is available. Even an approximation to the relationship constantly changes since it is dependent upon load conditions at the output and variations in the system itself due to wear, age and change in environmental conditions. The objective is to design a system that is insensitive to such changes.

To this end, "feedback" is applied to the open-loop system by sensing the value of the output and returning this information, possibly with modification, to the input for comparison with the input signal. Such a scheme is indicated in Fig. 5.1a where a feedback component is used to sense, modify and return the output signal of the open-loop system to the input terminals. Considering the open-loop system and the feedback component as two two-port devices, it is evident that ports 1 and 4 are in series with the input port 5 while ports 2 and 3 are in parallel with the output port 6. Such an interconnection is called the series-parallel feedback connection and this description is further substantiated by the form of the system graph in Fig. 5.1b. Ideally, the components should be unilateral with zero input admittance. However, this restriction is not necessary.

Given the state models of the open-loop system and the feedback component in Fig. 5.1a, the state model of the corresponding closed-loop system can be derived directly by combining the two state models and the interconnection equations from the indicated system graph.

As a practical example of feedback control, consider the electro-hydraulic system in Fig. 5.2. The output variable is the vertical displacement of the hydraulic ram piston. The position of the piston is to be precisely controlled by an input voltage v_{in} at the potentiometer FE.

The components enclosed by the dashed lines in Fig. 5.2 correspond to the open-loop system block in Fig. 5.1a. The amplifier-solenoid forms an ideal linear transducer. That is, the vertical displacement of point P is proportional to the voltage v_1 and further it is assumed that P undergoes a positive displacement (up) when v_1 is positive (A positive with respect to B). Under these assumptions, the state model of the open-loop system is obtained by combining the state model of the hydraulic subassembly given in (3.19) and the linear relationship for the amplifier-solenoid assembly: $\delta_1 = K_v v_1$. The result is

$$\frac{d}{dt}\begin{bmatrix} \delta_2 \\ \delta_2 \end{bmatrix} = \begin{bmatrix} A_h & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} \dot{\delta}_2 \\ \delta_2 \end{bmatrix} + \begin{bmatrix} K_h \cdot K_v & \frac{1}{M} \\ 0 & 0 \end{bmatrix}\begin{bmatrix} v_1 \\ f_2 \end{bmatrix} \tag{5.1}$$

$$i_1 = 0$$

where $K_h = \dfrac{K_{56}\,K_{21}}{M_6\,K_{22}}$, $M = M_6$ and A_h is defined by (3.20).

The feedback component in Fig. 5.1 consists of the potentiometer DE in Fig. 5.2 which furnishes a voltage v_4 proportional to the vertical position of the piston of the hydraulic ram. Its simple algebraic model is

$$\begin{aligned} v_4 &= K_f \delta_3 \\ f_3 &= 0 \end{aligned} \tag{5.2}$$

Finally, two independent variables are specified at the input and output terminals of the feedback system. These are

$$\begin{aligned} v_5 &= v_{in} \\ f_6 &= f_L \end{aligned} \tag{5.3}$$

Writing the interconnection equations from the system graph in Fig. 5.1b and combining them appropriately with (5.1-3), the state model of the feedback system takes the form

$$\frac{d}{dt}\begin{bmatrix} \dot{\delta}_2 \\ \delta_2 \end{bmatrix} = \begin{bmatrix} A_h & -K_v K_h K_f \\ 1 & 0 \end{bmatrix}\begin{bmatrix} \dot{\delta}_2 \\ \delta_2 \end{bmatrix} + \begin{bmatrix} K_v K_h & -\frac{1}{M} \\ 0 & 0 \end{bmatrix}\begin{bmatrix} v_{in} \\ f_L \end{bmatrix}$$

$$\delta_L = \delta_2 \tag{5.4}$$

In general, it is desirable to obtain from the mathematical model as much information as possible concerning the performance of the system without actually having to obtain a detailed solution. Assuming for the moment that the system is stable, the steady-state performance can be evaluated directly from the state model. Let a constant voltage input of magnitude V_1 be applied to the system simultaneously with a constant force loading at the output with magnitude F_L. Under steady-state conditions and constant excitation, the time derivatives of δ_2 vanish. Hence, from the first state equation of (5.4), we can write

$$0 = -K_v K_h K_f \delta_2 + K_v K_h V_1 - \frac{1}{M} F_L$$

or

$$\delta_L = \delta_2 = \frac{V_1}{K_f} - \frac{F_L}{K_v K_h K_f M} \qquad (5.5)$$

If the input potentiometer is calibrated so that numerically, V_1/K_f equals the desired value of δ_L, then

$$\delta_L = \delta_L - \frac{F_L}{K_v K_h K_f M} \qquad (5.6)$$

The effect of loading the piston shaft is clearly indicated by (5.6); the steady-state position error being zero only when $F_L = 0$.

The stability of the system can be obtained from the poles of the no-load transfer function. For simplicity, let the various constants in the state equations of (5.4) take on numerical values so as to give

$$\frac{d}{dt}\begin{bmatrix} \dot{\delta}_2 \\ \delta_2 \end{bmatrix} = \begin{bmatrix} -5 & -6 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} \dot{\delta}_2 \\ \delta_2 \end{bmatrix} + \begin{bmatrix} 6 & -2 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} v_{in} \\ f_L \end{bmatrix} \qquad (5.7)$$

Following the procedure outlined in Section 4 for obtaining the transfer function from the state model, we set the through variable $f_L = 0$ in (5.7), assume zero initial conditions for the state variables $\dot{\delta}_2$ and δ_2 and take the Laplace transform of both sides to obtain

$$s\begin{bmatrix} \dot{\delta}_2(s) \\ \delta_2(s) \end{bmatrix} = \begin{bmatrix} -5 & -6 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} \dot{\delta}_2(s) \\ \delta_2(s) \end{bmatrix} + \begin{bmatrix} 6 \\ 0 \end{bmatrix} v_{in}(s)$$

or

$$\begin{bmatrix} s+5 & 6 \\ -1 & s \end{bmatrix}\begin{bmatrix} \dot{\delta}_2(s) \\ \delta_2(s) \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix} v_{in}(s)$$

Solving for the state vector gives

$$\begin{bmatrix} \dot{\delta}_2(s) \\ \delta_2(s) \end{bmatrix} = \frac{1}{s(s+5)+6}\begin{bmatrix} s & -6 \\ 1 & s+5 \end{bmatrix}\begin{bmatrix} 6 \\ 0 \end{bmatrix} v_{in}(s) \qquad (5.8)$$

Combining the last equation of (5.8) with the s-domain terminal relation $\delta_L(s) = \delta_2(s)$, we obtain the no-load transfer function

$$G(s) = \frac{\delta_L(s)}{v_{in}(s)} = \frac{\delta_2(s)}{v_{in}(s)} = \frac{6}{s^2 + 5s + 6} \tag{5.9}$$

or

$$\delta_L(s) = \frac{6}{(s+2)(s+3)} \cdot v_{in}(s) \tag{5.10}$$

The poles of $G(s)$ are $s_1 = -2$ and $s_2 = -3$. Thus, if the input variable v_{in} is bounded, the time function $\delta_L(t)$ corresponding to $\delta_L(s)$ in (5.10) is a stable time function and the feedback system is stable. Note that we have assumed no force loading at the output.

When the components are modeled as ideal transducers, a very useful relationship exists between the no-load transfer function $G(s)$ of a closed-loop system and the no-load transfer functions of the open-loop system $G_o(s)$ and the feedback component $G_f(s)$. With regard to the system graph of Fig. 5.1b, we have the following relations holding between the s-domain across variables.

$$\begin{aligned} x_1(s) &= x_5(s) - x_4(s) \\ x_6(s) &= x_3(s) = x_2(s) \end{aligned} \tag{5.11}$$

By definition

$$G_o(s) = \frac{x_2(s)}{x_1(s)}, \quad G_f(s) = \frac{x_4(s)}{x_3(s)} \tag{5.12}$$

Combining (5.11) and (5.12) we obtain

$$G(s) = \frac{x_6(s)}{x_5(s)} = \frac{G_o(s)}{1 + G_o(s)\,G_f(s)} \tag{5.13}$$

Thus, for unilateral components, the transfer function of the closed-loop system is easily expressed in terms of the transfer functions of the open-loop system and the feedback component. Since the poles of $G(s)$ determine stability and transient response characteristics of a closed-loop system, much emphasis is placed on the zeros of the expression $1 + G_o(s)G_f(s)$. Essentially, all linear feedback system design techniques such as root-locus and Nyquist plots are based on this expression.

The complete transient response of a feedback system for a given input can only be obtained by a full solution of the state model. Such a solution can be accomplished either by computer using a numerical technique or by a hand calculation using Laplace transforms or functions of matrices. For simple problems, hand calculation suffices. As a final example, we consider a complete solution to the state model of (5.7) with Laplace transforms.

Let the initial conditions on the state vector be taken as

$$Z_o = \begin{bmatrix} \dot{\delta}_2(0) \\ \delta_2(0) \end{bmatrix} = \begin{bmatrix} 0 \\ -2 \end{bmatrix} \tag{5.14}$$

and let

$$
\begin{aligned}
v_{in} &= V_1\, u(t) \\
f_L &= F_L\, u(t)
\end{aligned}
\qquad (5.15)
$$

define the independent variables where u(t) is the unit step function

$$
\begin{aligned}
u(t) &= 1 \text{ for } t \geq 0 \\
&= 0 \text{ otherwise}
\end{aligned}
\qquad (5.16)
$$

with Laplace transform

$$
L[u(t)] = \frac{1}{s} \qquad (5.17)
$$

Taking the Laplace transform of both sides of (5.7) and applying (5.14-17), we obtain

$$
s\begin{bmatrix} \dot{\delta}_2(s) \\ \delta_2(s) \end{bmatrix} = \begin{bmatrix} -5 & -6 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} \dot{\delta}_2(s) \\ \delta_2(s) \end{bmatrix} + \begin{bmatrix} 0 \\ -2 \end{bmatrix} + \begin{bmatrix} 6 & -2 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} V_1/s \\ F_L/s \end{bmatrix}
\qquad (5.18)
$$

Solving for the s-domain state variables gives

$$
\begin{bmatrix} \dot{\delta}_2(s) \\ \delta_2(s) \end{bmatrix} = \frac{1}{(s+2)(s+3)}\begin{bmatrix} s & -6 \\ 1 & s+5 \end{bmatrix}\begin{bmatrix} 0 \\ -2 \end{bmatrix} + \frac{1}{(s+2)(s+3)}\begin{bmatrix} s & -6 \\ 1 & s+5 \end{bmatrix}\begin{bmatrix} 6 & -2 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} V_1/s \\ F_L/s \end{bmatrix}
\qquad (5.19)
$$

The form of (5.19) can be compared directly with (4.11). A partial fraction expansion of (5.19) has the form

$$
\begin{aligned}
\begin{bmatrix} \dot{\delta}_2(s) \\ \delta_2(s) \end{bmatrix} &= \frac{R_1}{(s-s_1)} + \frac{R_2}{(s-s_2)} + \frac{F_1}{(s-s_1)} + \frac{F_2}{(s-s_2)} + \frac{F_a}{(s-s_a)} \\
&= \frac{1}{s+2}\begin{bmatrix} 12 \\ -6 \end{bmatrix} + \frac{1}{s+3}\begin{bmatrix} -12 \\ 4 \end{bmatrix} + \frac{1}{s+2}\begin{bmatrix} 6V_1 & -2F_L \\ -3V_1 & +F_L \end{bmatrix} + \\
&\quad + \frac{1}{s+3}\begin{bmatrix} -6V_1 & +2F_L \\ 2V_1 & -\frac{2}{3}F_L \end{bmatrix} + \frac{1}{s}\begin{bmatrix} 0 \\ V_1 \quad -\frac{1}{3}F_L \end{bmatrix}
\end{aligned}
\qquad (5.20)
$$

where, for example, the vector

$$F_2 = \lim_{s \to -3} \frac{(s+3)}{(s+2)(s+3)} \begin{bmatrix} s & -6 \\ 1 & s+5 \end{bmatrix} \begin{bmatrix} 6 & -2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1/s \\ F_L/s \end{bmatrix}$$

$$= \lim_{s \to -3} \frac{1}{(s+2)(s)} \begin{bmatrix} 6s & -2s \\ 6 & -2 \end{bmatrix} \begin{bmatrix} V_1 \\ F_L \end{bmatrix} = \begin{bmatrix} -6V_1 + 2F_L \\ 2V_1 - \frac{2}{3}F_L \end{bmatrix}$$

Finally, applying the transform pair

$$L[A\, e^{at}] = \frac{A}{s + a}$$

we obtain as the inverse Laplace transform of (5.20) the time solution

$$\begin{bmatrix} \dot{\delta}_2(t) \\ \delta_2(t) \end{bmatrix} = \begin{bmatrix} \dot{\delta}_L(t) \\ \delta_L(t) \end{bmatrix} = \begin{bmatrix} 12 & +6V_1 & -2F_L \\ -6 & -3V_1 & +F_L \end{bmatrix} e^{-2t} +$$

$$+ \begin{bmatrix} -12 & -6V_1 & +2F_L \\ 4 & +2V_1 & -\frac{2}{3}F_L \end{bmatrix} e^{-3t} + \begin{bmatrix} 0 \\ V_1 - \frac{1}{3}F_L \end{bmatrix} \tag{5.21}$$

Assuming specific magnitudes of the input voltage and output load force $V_1 = 0.5$ and $F_L = 3$, the time solution becomes

$$\begin{bmatrix} \dot{\delta}_L(t) \\ \delta_L(t) \end{bmatrix} = \begin{bmatrix} 9 \\ -4.5 \end{bmatrix} e^{-2t} + \begin{bmatrix} -9 \\ 3 \end{bmatrix} e^{-3t} + \begin{bmatrix} 0 \\ -0.5 \end{bmatrix} \tag{5.22}$$

Notice that the initial conditions (5.14) are satisfied and the solution is stable. As $t \to \infty$, $\delta_L \to -0.5$, resulting in a steady-state error of 1 unit. This error is caused by the output load force. With $F_L = 0$, the solution (5.21) reduces to

$$\begin{bmatrix} \dot{\delta}_L(t) \\ \delta_L(t) \end{bmatrix} = \begin{bmatrix} 15 \\ -7.5 \end{bmatrix} e^{-2t} \begin{bmatrix} -15 \\ 5 \end{bmatrix} e^{-3t} + \begin{bmatrix} 0 \\ 0.5 \end{bmatrix} \tag{5.23}$$

Thus, as predicted by (5.6), the steady-state error between input and output is zero when the load force is zero.

REFERENCES

Feedback Control Systems:

Chestnut, H. and Mayer, R. W., Servomechanisms and Regulating System Design, Volume 1, John Wiley and Sons, Inc., 1951.

Kuo, B. C., Automatic Control Systems, Prentice-Hall Inc., 1962.

Thaler, G. J. and Brown, R. G., Analysis and Design of Feedback Control Systems, second edition, McGraw-Hill Book Co., Inc., 1960.

Truxal, J. G., Automatic Feedback Control System Synthesis, McGraw-Hill Book Co., Inc., 1955.

Linear Graph Theory:

Seshu, S. and Reed, M. B., Linear Graphs and Electrical Networks, Addison-Wesley Pub. Co., Inc., 1961.

Koenig, H. E. and Frame, J. S., An Introduction to Functions of Matrices with Applications, Part III, IEEE Spectrum, to appear, Spring, 1964.

Mathematical Models of Physical Systems:

Koenig, H. E. and Blackwell, W. A., Electromechanical System Theory, McGraw-Hill Book Co., Inc., 1961.

Koenig, H. E., Tokad, Y. and Kesavan, H. K., Analysis of Discrete Physical Systems, McGraw-Hill Book Co., Inc., in publication.

Matrix Theory:

Basic Theory:

Hohn, F. E., Elementary Matrix Algebra, MacMillan Co., 1958.

Perlis, S., Theory of Matrices, Addison-Wesley Pub. Co., Inc., 1952.

Advanced Theory and Functions of Matrices:

Frame, J. S., An Introduction to Functions of Matrices with Applications, Parts I and II, IEEE Spectrum, to appear, Spring, 1964.

Gantmacher, F. R., The Theory of Matrices, Vol. I, Chelsea Pub. Co., 1960.

Transform Methods and Complex Variables

Aseltine, J. A., Transform Method in Linear System Analysis, McGraw-Hill Book Co., Inc., 1958.

Churchill, R. V., Complex Variables and Applications, second edition, McGraw-Hill Book Co., Inc., 1960.

Lepage, W. R., Complex Variables and the Laplace Transform for Engineers, McGraw-Hill Book Co., Inc., 1961.

Process Control in the Ceramic Industry

Ernest Hucke

ABSTRACT

The basic components in a block diagram depicting a general closed loop process control system are covered. Emphasis is given to pneumatic and electronic sensors, transducers, transmitters, as well as controllers. Typical control systems, using some of these components are shown in block diagram, and their characteristics are discussed briefly.

I. INTRODUCTION

The importance of representing each unit of control equipment as a block and to show all units inthe correct relative position for easier evaluation of system performance has been discussed. In order to properly construct a complete system, however, one must know something about the components which will make up the control system.

A simple closed loop control system is shown in Fig. 1.

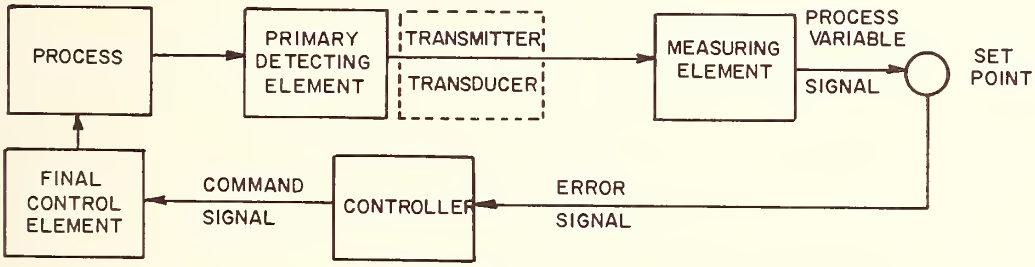


Fig. 1 Block Diagram of Closed Loop Control System

The first blocks that are encountered and discussed in the basic control loop are the Primary Detecting Element and the Measuring Element. It is generally accepted that the primary detecting element is that portion of the feedback that either utilizes or transforms energy from the controlled medium to produce a signal which is a function of the value of the directly controlled variable. The measuring element is that portion of the feedback elements which converts the signal from the primary detecting element to a form comparable with the set point reference input. This chapter will be limited to hardware that is being used in the Ceramic Industry and which can form part of a control loop.

II. Temperature Detecting Elements

Filled Thermal Systems

Filled-system thermometers are generally classified as vapor, gas or mercury filled, depending uponthe fluid in the sealed measuring unit. The measuring element is connected by capillary tubing to the primary detecting element, in this case, a bulb thermally coupled to the process. The measuring element converts the motion proportional to temperature to a form compatible with the reference input, and, in most cases, is also mechanically linked to a pen or pointer or to a control element.

The response speed of a typical thermometer-bulb is indicated in Figure 2, and a thermal system is shown in Figure 3.

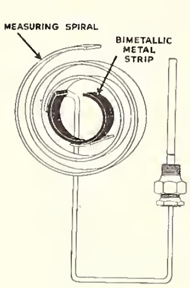


Fig. 3 Thermal System with Case Compensation

Thermocouples

Two dissimilar metals welded together at one end and heated at this junction will develop a voltage at the free ends. This voltage is proportional to the temperature difference between the welded and the free end. Such a device can be used with a voltage measuring galvanometer to construct a pyrometer as shown in Figure 4.

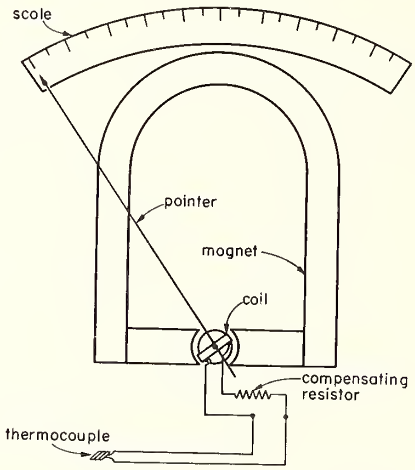


Fig 4. Pyrometer Using Simple Deflection Galvanometer

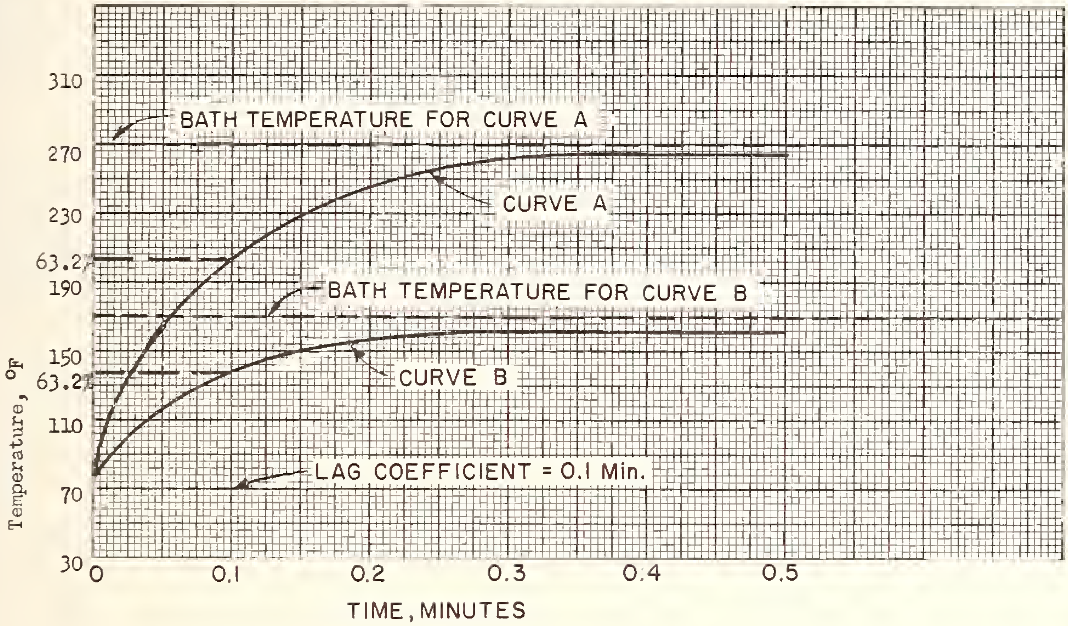


Fig. 2 Speed of Response of Thermometer Bulb

By proper selection of the two metals, temperature ranges of from -300°F to +4000°F can be covered by thermocouples. Figure 5 shows typical calibration curves.

Radiation

A number of small thermocouples arranged in series is called a Thermopile. In normal use infrared radiation from a heated body is focused by a lens onto the measuring junctions of these thermocouples. The radiation detector is not immersed in the heated source as with other temperature measuring devices. The thermopile can be used for temperature ranges of from ambient to 7000°F.

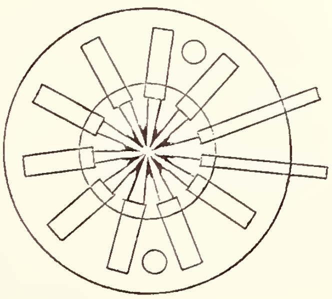


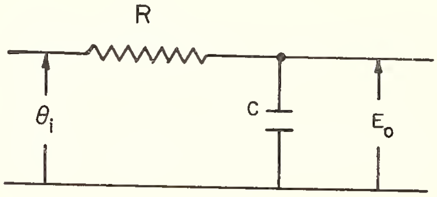
Fig. 6 Radiation Pyrometer-Thermopile

Resistance Thermometers

The property of metals to increase in electrical resistance as temperatures rises provides a temperature measurement known as resistance thermometry. The primary detecting element is a wire wound resistor, called a resistance thermometer bulb. The bulb is connected to a measuring instrument incorporating a Wheatstone bridge. The speed of response of a typical resistance thermometer is given in Figure 7.

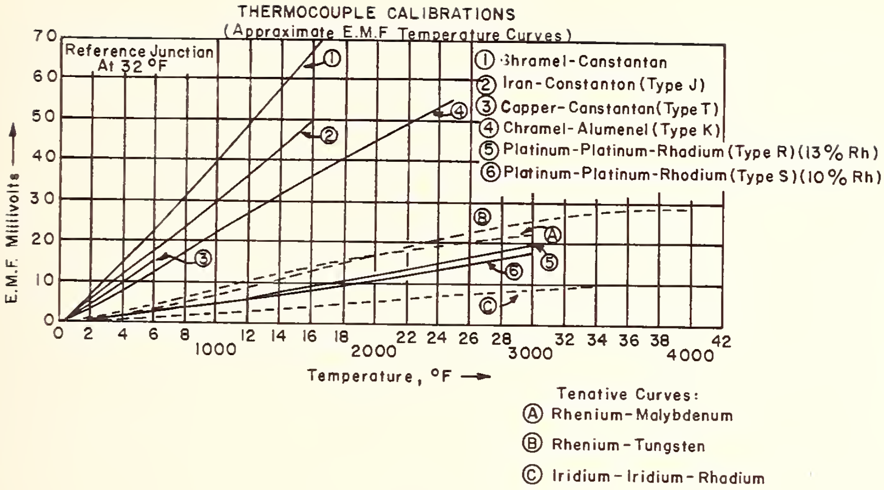
Response of Temperature Sensors

The dynamic response of most temperature sensors is determined by a thermal resistance and a thermal capacitance in the signal flow path. A bare thermocouple has a significant thermal resistance at the boundary between the environment and the junction. Such an arrangement is analogous to an electrical R-C filter circuit as shown in Figure 8. Its transfer function is a first order linear lag, described by the equation [1].



$$G(s) = \frac{E_o(s)}{\theta_i(s)} = \frac{1}{RC + 1} \qquad [1]$$

Fig. 8 An Electrical R-C Filter Circuit



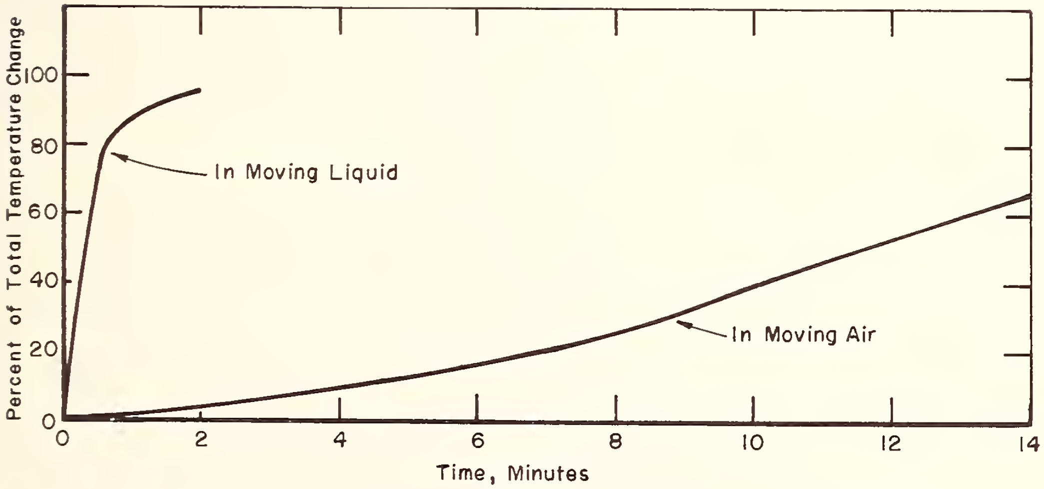


Fig. 7 Speed of Response of Resistance Thermometer Bulb

where:

R = thermal (film) resistance
C = thermal capacitance
RC = T = time constant
θ_i = temperature input
E_o = EMF output
s = Laplace operator
G(s) = transfer function

A protective well surrounding the thermocouple will add another section in cascade to the network described above. Since there is a loading or interacting effect, the overall transfer function will be of the following form:

$$G(s) = \frac{1}{R_1R_2C_1C_2s^2 + (R_1C_1 + R_2C_2 + R_1C_2)\, s + 1} \qquad [2]$$

The equation is factorable into the form

$$G(s) = \frac{1}{(T_a s + 1)\,(T_b s + 1)} \qquad [3]$$

Where R_1, C_1 and R_2, C_2 refer, respectively, to the well and to the junction; the time constants T_a, T_b are the roots of the denominator quadratic (2).

Radiation pyrometers, bare-bulb resistance thermometers and bare-bulb filled systems exhibit an approximate first order response, however, their time constants vary considerably with the various media and arrangements in question (1), (3), (5). Figures 8, 9, and 10 give time responses of typical temperature measuring devices.

Pyrometric Cones

They are cones made on the form of slender, truncated, trigonal pyramids from raw materials used in the manufacture of many ceramic products. Exposed to the same firing and furnace conditions as the product, the physical and chemical changes in the cones are indicative as to the changes occurring in the product as a function of temperature and time (8).

III. Pressure Detecting Elements

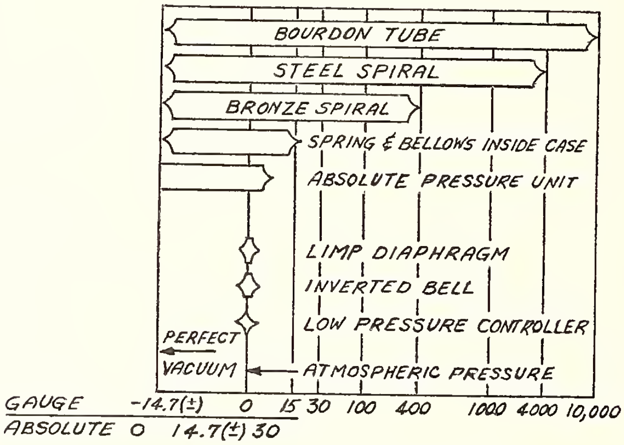


Fig. 11 Application Chart of Pressure Measuring Devices

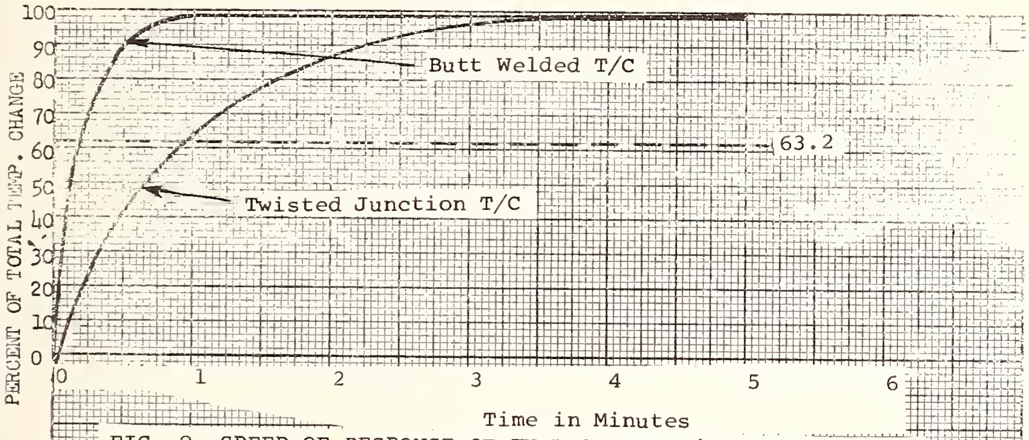


FIG. 8 SPEED OF RESPONSE OF THERMOCOUPLE (IN STILL LIQUID)

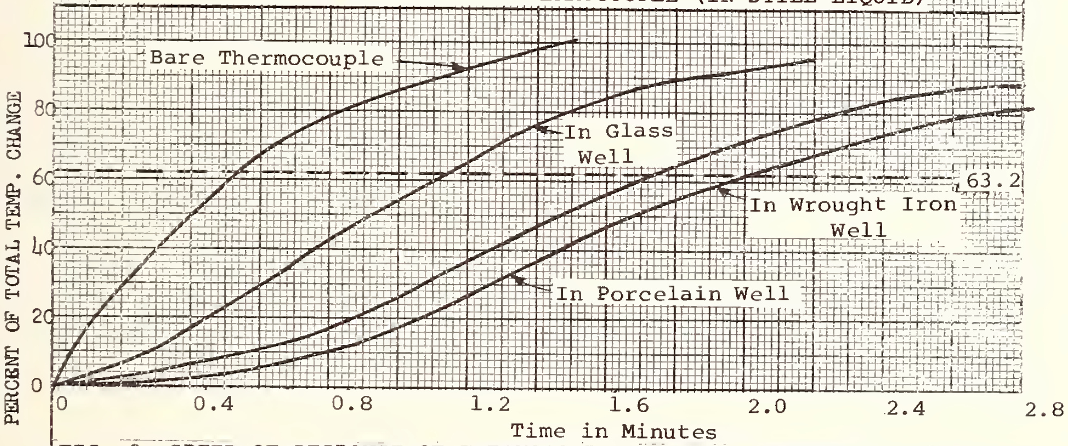


FIG. 9 SPEED OF RESPONSE OF THERMOCOUPLE, WITH & WITHOUT PROTECTING WELLS (IN AIR)

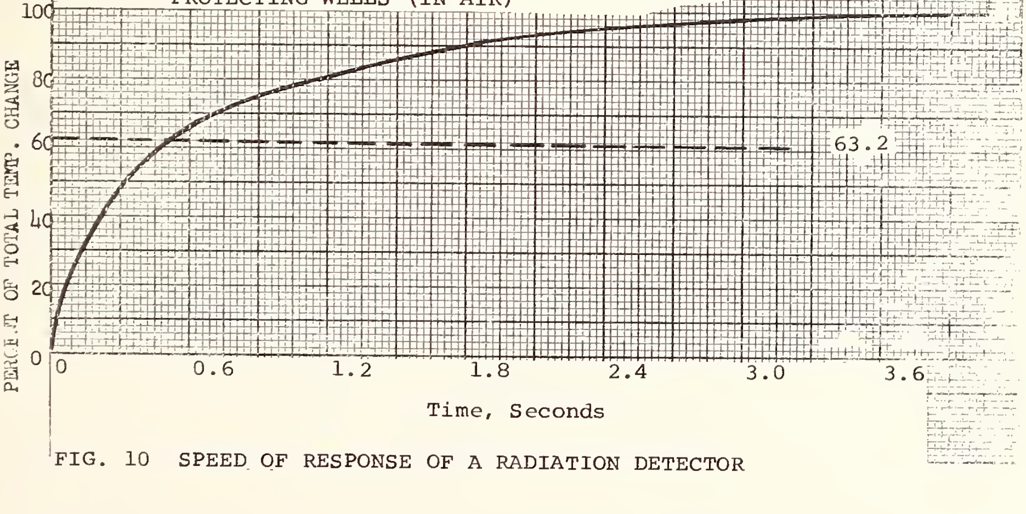


FIG. 10 SPEED OF RESPONSE OF A RADIATION DETECTOR

There are 3 main types of elastic pressure elements; diaphragms, bellows and Bourdon tubes. All follow "Hooks Law", which states that stress is proportional to strain, or deflection is proportional to the pressure applied, Mathematically

$$x = \frac{F}{K} \qquad [4]$$

and

$$F = P_c A \qquad [5]$$

where
x = output deflection
F = force
k = spring gradient
P_c = inside pressure
A = effective area of the sensor

In all three devices, the pressure built-up in the enclosure is determined by the capillary restriction and the volume of the capsule. For small changes in pressure it is possible to define a pneumatic resistance R and a pneumatic capacitance C. Then the resulting transfer function relating capsule pressure ΔP_c to input pressure ΔP_i is given by

$$G(s) = \frac{\Delta P_c(s)}{\Delta P_i(s)} = \frac{1}{RCs + 1} \qquad [6]$$

Combining equations 4, 5, and 6 yields

$$G(s) = \frac{\Delta X(s)}{\Delta P_i(s)} = \frac{A/k}{RCs + 1} \qquad [7]$$

Equation 7 can be represented by the block diagram shown below

$$\Delta P_i(s) \longrightarrow \boxed{\frac{1}{RCs + 1}} \xrightarrow{\Delta P_c(s)} \boxed{A} \xrightarrow{\Delta F(s)} \boxed{\frac{1}{k}} \longrightarrow \Delta X(s)$$

In some instances, the attached mass will become significant; this would lead to a third order transfer function.

Diaphragm - Non-Metallic

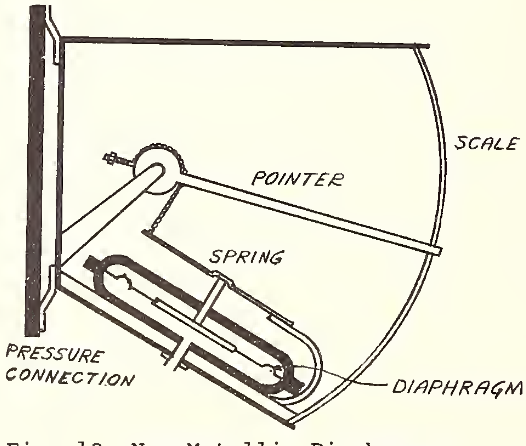


Fig. 12 Non Metallic Diaphragm

One type of diaphragm is made of flexible leather, (oiled to assure flexibility over a long period of time), colon leather, or of noeprene. The diaphragm is generally opposed by a spring, which is deflected by an amount directly proportional to the applied pressure. A linkage connects the diaphragm to the pen and/or measuring element. A typical structure is shown in Figure 12.

Metallic Diaphragm

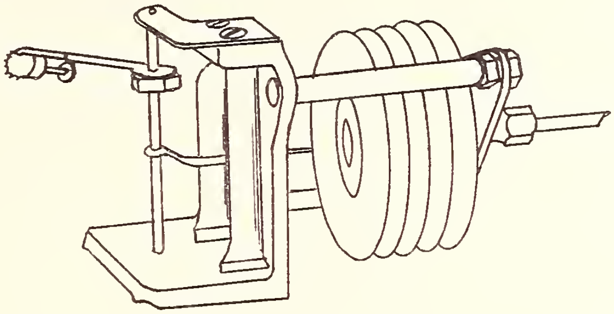


Fig. 13 Metallic Diaphragm

Flexible metal diaphragms are also used and their action is similar to that previously described. The diaphragm, however, is of thin brass or other flexible metal chosen to be compatible with the corrosive action of the process.

Bellows Elements

The bellows element is axially flexible and is a one piece expansible and collapsible unit. It is usually formed of thin seamless tubing made into a corrugated unit either hydraulically or mechanically. These elements are made out of phosphor bronze, brass, stainless steel, beryllium copper and other non corrosive materials as required by the corrosive condition of the process. They are calibrated for various pressure ranges by a bellows spring, which is external, and therefore not subject to corrosion by the fluid under measurement. For vacuum and compound ranges an internal spring is also required as the bellows materials itself is not suitable for accurate calibration.

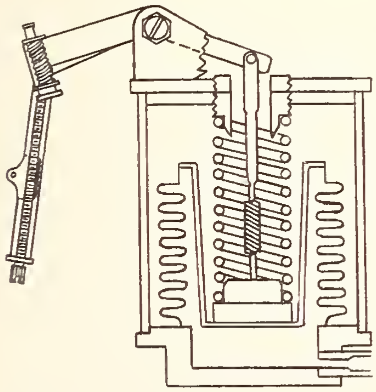


Fig. 14 Spring and Bellows Assembly for Pressure Measurement

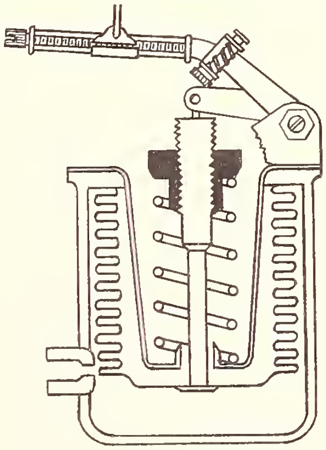


Fig. 15 Spring and Bellows Assembly for Vacuum Measurement

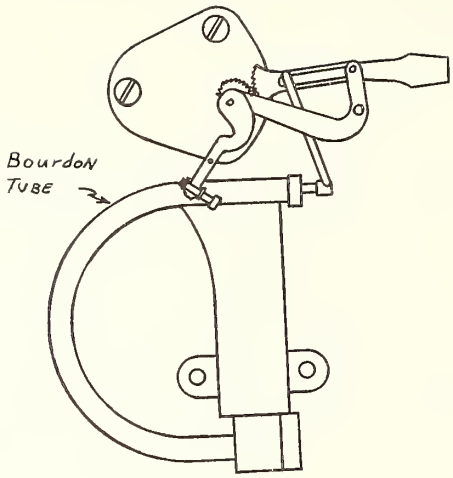


Fig. 16 Simple Application of Bourdon Tube

The Bourdon tube is probably the oldest pressure detecting element and is the most widely used pressure measuring device in industry. The tube itself is flattened on diametrically opposite sides to produce an approximately elliptical or oval section and is then bent into an arc of a circle. Pressure applied to the inside of the tube will distend the flattened positions and tend to restore the original circular cross section. This change in shape tends to straighten the tube and in so doing moves the tip a sufficient distance to actuate the linkage. The most commonly used Bourdon tubes are made out of phosphor bronze, alloy steel, stainless steel and beryllium copper. Their ranges cover from vacuum to 10,000 psi. Bored out of solid steel forgings they will detect pressure up to 80,000 psi.

The Spiral

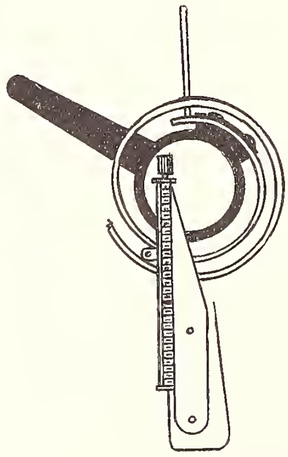


Fig. 17 A Spiral Gauge

Another useful gauge is the spiral gauge which is formed out of a metal tube, flattened in cross section and wound into a spiral shape. The open inner end of the spiral is connected to the pressure source while the linkage is attached to the closed outer end. When pressure is applied, the spiral, which is firmly fastened at its inner end, unwinds, causing the outer end to rotate in a counter-clockwise rotation.

The Inverted Bells

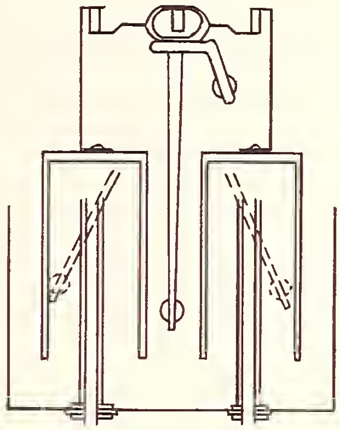


Fig. 18 An Inverted Bell Pressure Gauge

Another pressure measuring device is the inverted bell. Two inverted bells are shown detecting the differential pressure. The pressures are being piped to the bells on opposite ends of a balanced beam from which the bells are suspended. Static pressure can be detected by a single bell with a counterweight suspended from the other end of the beam. The dynamic response of the inverted bell is second order.

Electrical Pressure Transducers

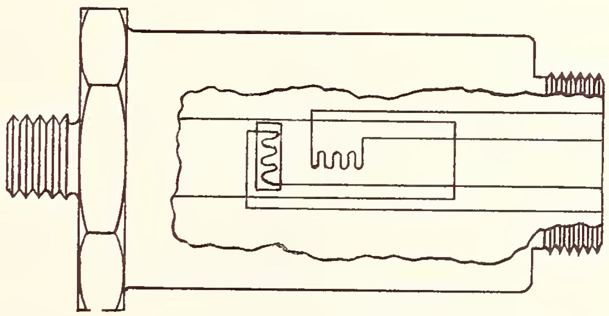


Fig. 19 An Electrical Pressure Transducer

The strain gage is a common pressure-to-electrical voltage transducer. The strain gage itself consists of wire grid bonded to the surface of a small slip of rag bond or bakelite impregnated paper. When the gages are cemented to a surface which is subjected to load, the wires, constituting the grids are stretched, thereby changing their diameter and length and consequently their electrical resistance. Through measurement of the change in resistance of the wire grid with the gage under stress, an evaluation of the force producing the load is obtained. The strain gages are electrically connected to form two legs of a Wheatstone bridge circuit. The stress to voltage response is instantaneous.

IV. FLOW METERS

Of the several types of flowmeters, the most frequently used is the differential pressure or heat meter, which measures the differential pressure produced by a restriction in the flow line. This resitriction, which is the primary element, can be produced in a number of ways. Common devices for measuring flow in this way are orifice plate, the venturi tube, the flow nozzle and the pitot tube shown respectively in Figures 20 through 23.

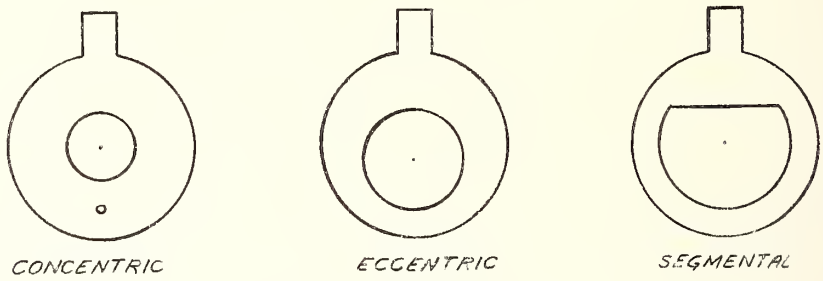


Fig. 20 Orifice Plates

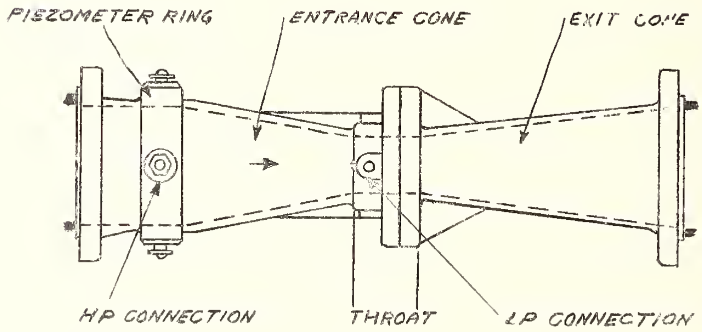


Fig. 21 A Venturi Tube

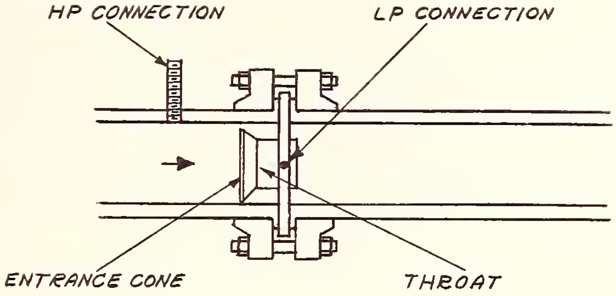


Fig. 22 A Flow Nozzle

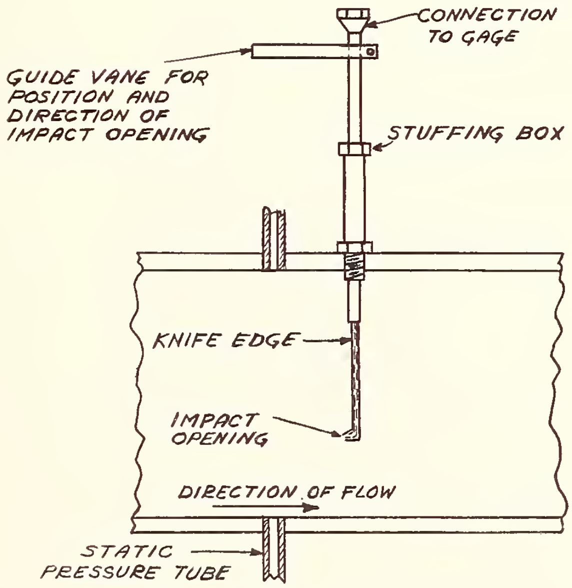


Fig. 23 A Pitot Tube

Principle of Operation

There are many types of measuring elements, which measure the relationship of flow rate and differential pressure where the flow varies as the square root of the differential pressure of Q = K H

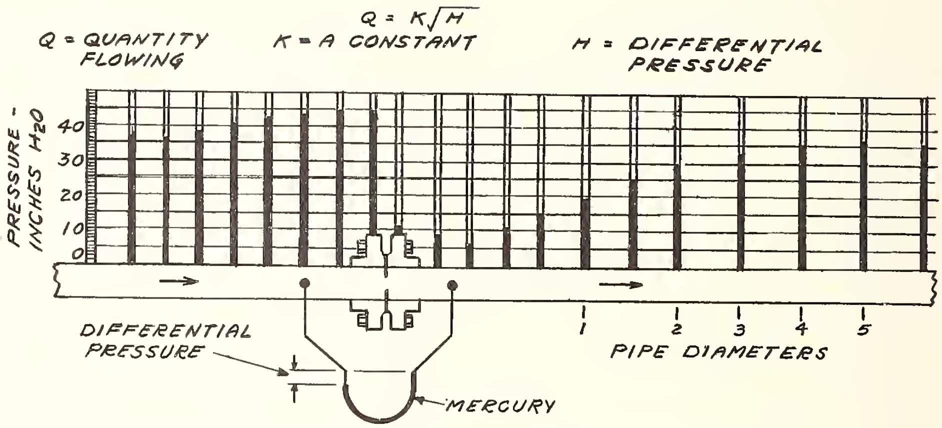


Fig. 24 Differential Pressure Across an Orifice

The Mercury Float Type Flow Meter

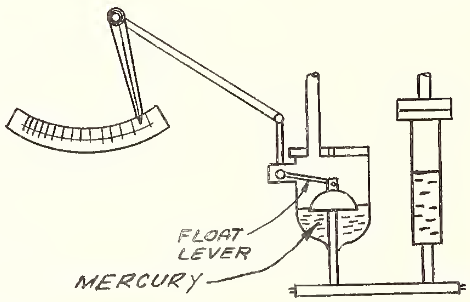


Fig. 25 Schematic of Mechanical Flow Meter

This device basically comprises a mercury filled manometer with a large chamber containing the float and the range tube in a smaller chamber.

The Bell Type

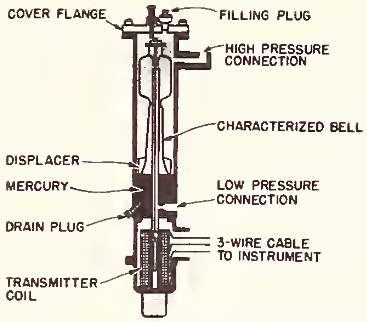


Fig. 26 Ledoux Bell Meter Body

The bell type flow meter is essentially a pressure chamber, which contains a characterized bell floating in a pool of mercury. The low pressure is applied to the inside of the bell and the high pressure to the outside. This type sensor is basically of the 2nd order (no specific data are available).

The Bellows Type

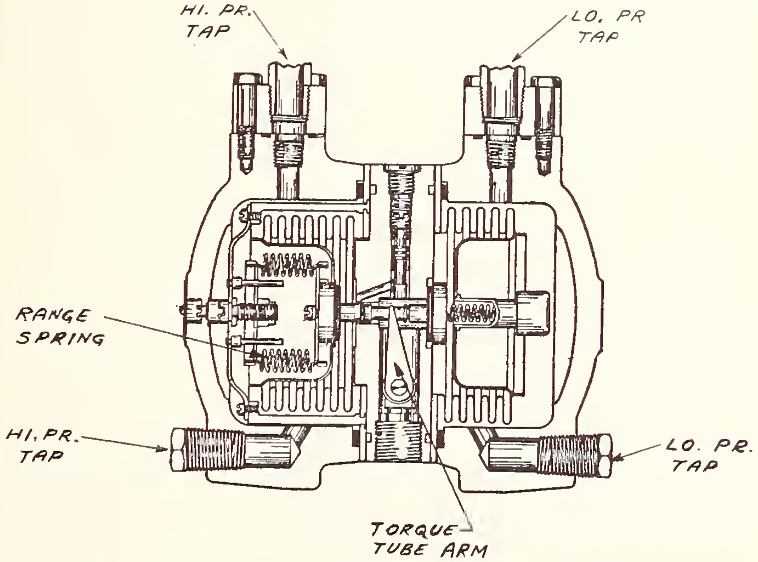


Fig. 27 Differential Pressure Converter, Bellows Type

The differential pressure acts against metal bellows, with a torque tube to transmit the bellows motion. The liquid filled damping system provides essentially first order response as the typical curves of Figure 28 indicate.

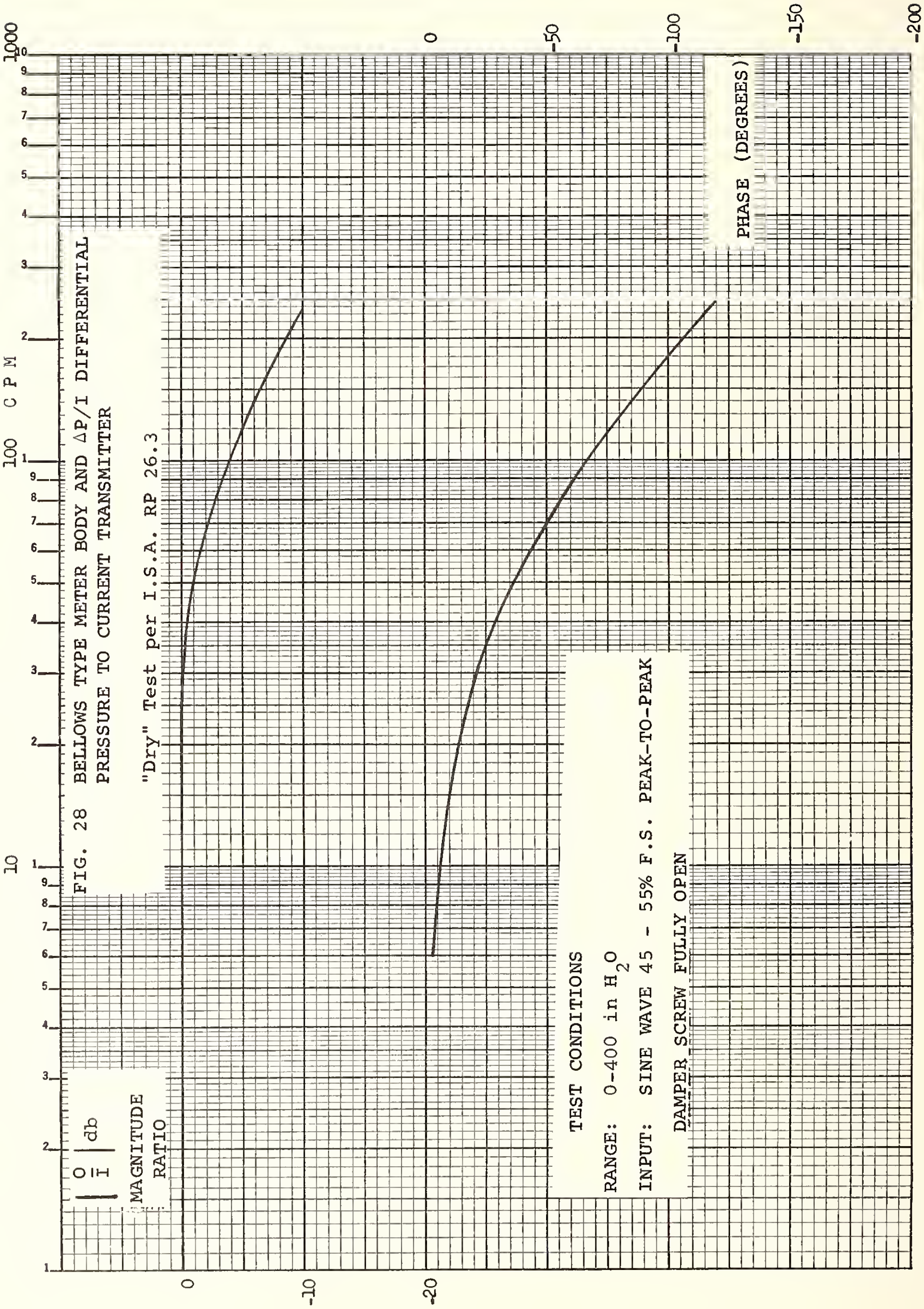


FIG. 28 BELLOWS TYPE METER BODY AND ΔP/I DIFFERENTIAL PRESSURE TO CURRENT TRANSMITTER

"Dry" Test per I.S.A. RP 26.3

The Force-Balance Type

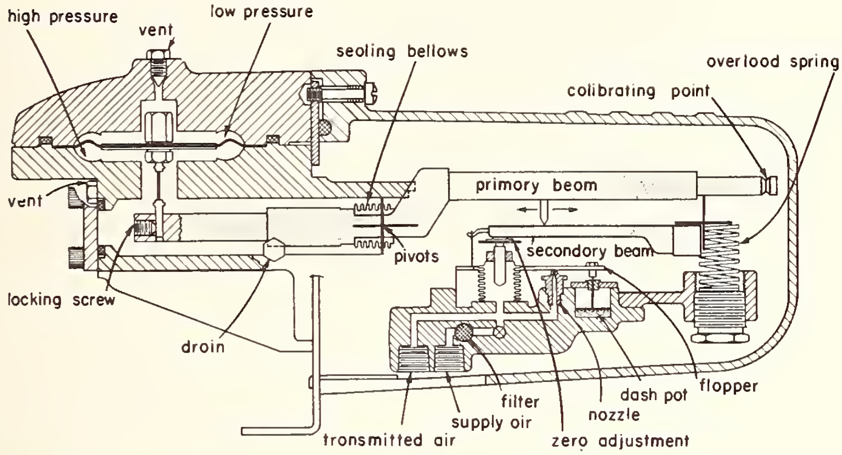


Fig. 29 Differential Pressure Converts, Force-Balance Type

In this type device the pressure differential is applied across a diaphragm. By mechanical and pneumatic means the force on the diaphragm is balanced so that the output pressure is proportional to the applied differential. The cascading of mechanical and pneumatic systems results in a high order transfer function, as shown for a typical device in Figure 30.

The Area Type Meter

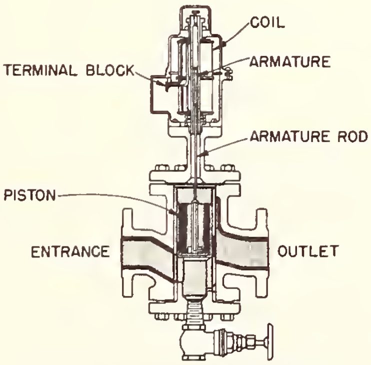
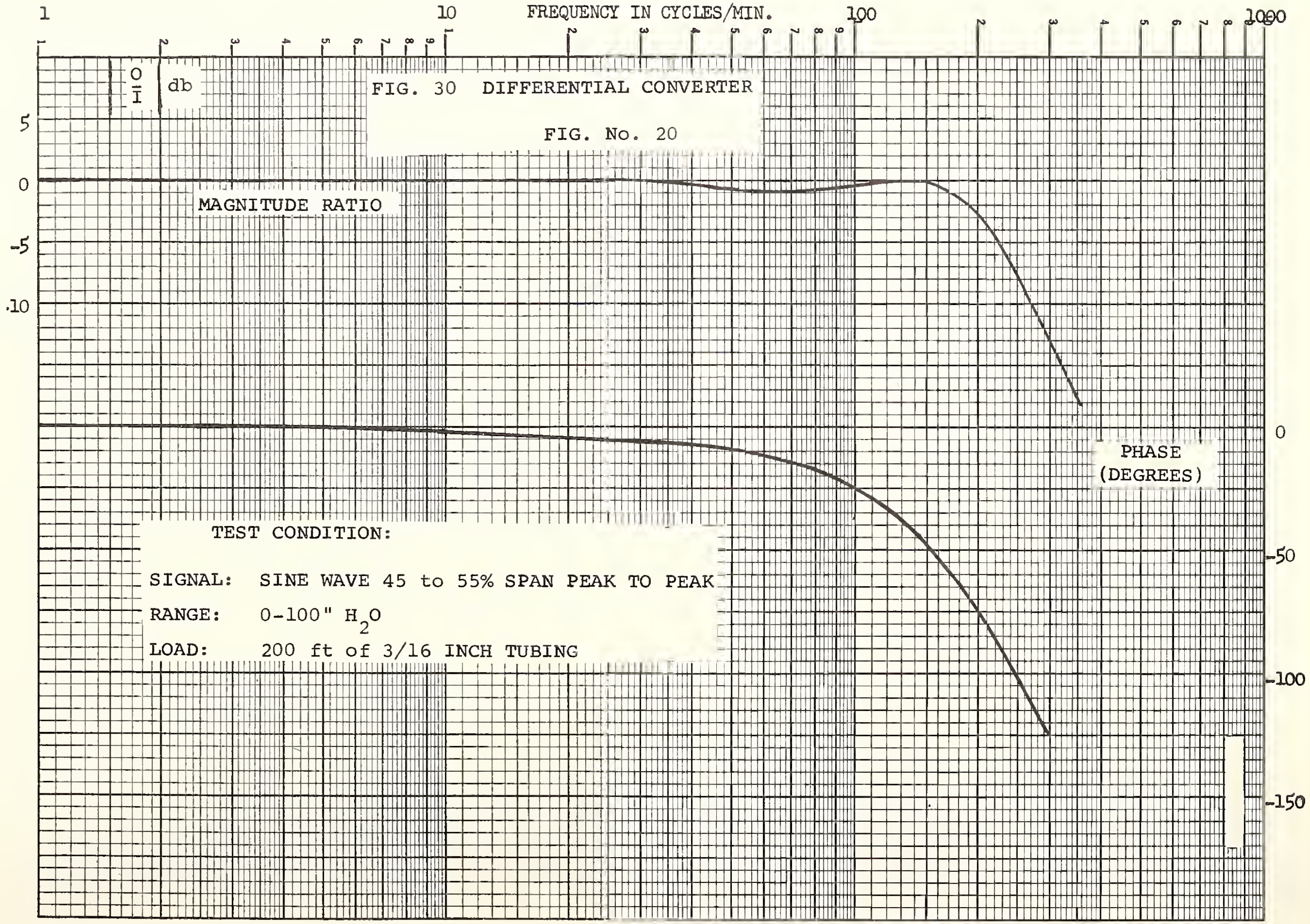


Fig. 31 Area Meter Body

In the area type meter, the area of the orifice is varied to maintain a constant differential pressure. The flow enters the meter body horizontally, is deflected upwards against a floating piston, and passes downstream through an orifice on the side opposite the inlet. A passageway is provided from the low-pressure side of the meter to the top of the piston, so that the downstream pressure always exists on top of the piston. As the piston moves up on an increase of flow, it opens the orifice just enough to maintain a constant pressure differential across the high and low-pressure sides of the orifice. Referring to Basic formula, Q = A H, since the area is varied and not the differential pressure, the flow will vary directly as the change in area. The flow is indicated by the position of the piston. Electric, electronic or pneumatic position transmission is available. The response is of the second order or higher.



1
10
FREQUENCY IN CYCLES/MIN.
100
1000
O/I db
5
0
-5
-10
FIG. 30 DIFFERENTIAL CONVERTER
FIG. No. 20
MAGNITUDE RATIO
PHASE (DEGREES)
0
-50
-100
-150
TEST CONDITION:
SIGNAL: SINE WAVE 45 to 55% SPAN PEAK TO PEAK
RANGE: 0-100" H_2O
LOAD: 200 ft of 3/16 INCH TUBING

The Rotameter

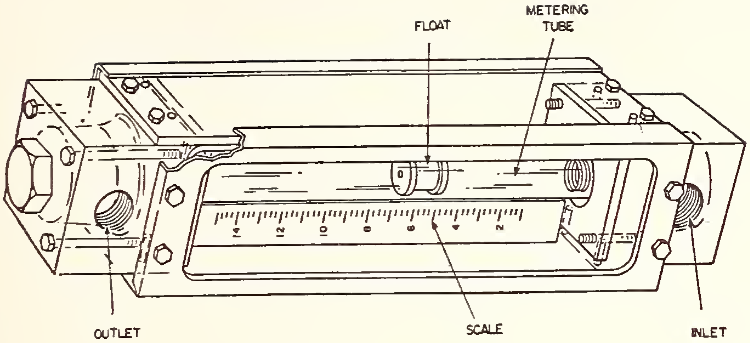


Fig. 32 Rotameter

The rotameter consists of an upright tapered tube, large end up, containing a weighted plummet, which is lifted to the position of equilibrium between the downward force of the plummet and the upward force of the fluid flowing past it through an annular orifice. The transfer function is of the second order since mass, spring and damping are present.

Magnetic Flow Sensor

Based on Faraday's law of electro-magnetic induction, the poles of a permanent or electro magnet are arranged so that the flux lines between North and South pole transverse the flow at right angles. The fluid flowing acts as a conductor moving through the magnetic field and causes an emf to be induced in the fluid which is proportional to velocity. Electrodes flush with the inner pipe surface in contact with the fluid measure this EMF. The response is practically instantaneous with the change of flow.

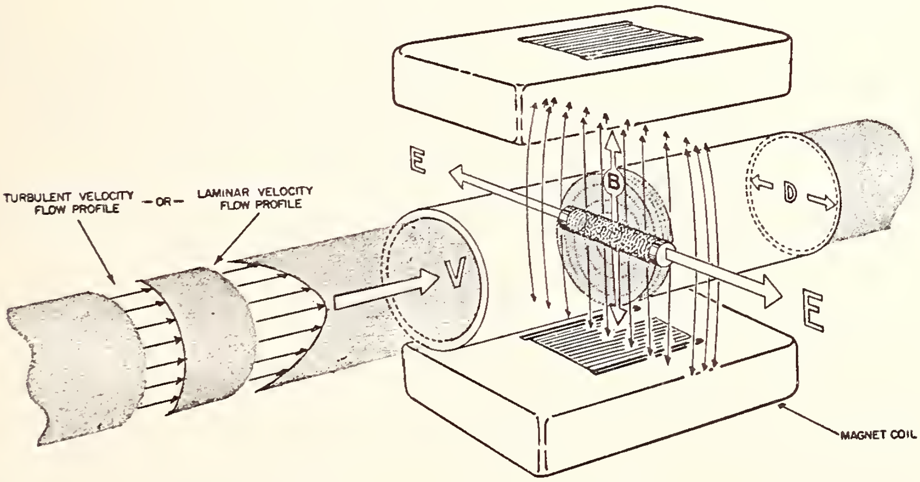


Fig. 33 Principle of Magnetic Flow Meter

Mass Flow-Dry Solids

The radioactive density meter frequently used to measure % solids in slurries or % concentration of acids has been found to be quite accurate in measuring mass flow per unit time of dry solids. The response is instantaneous.

Pioneering applications are in the cement and gypsum industries. The important conditions are:

1) Constant velocity

This is most often obtained by using free fall of the material at a point within 1-1/2 ft. of the start of free fall.

2) No large changes in composition or moisture

Use of steam below point of measurement may change moisture content of material.

3) Sufficient quantity of material flow to absorb minimum amount of radiation for instrument response. (1#/ft^3 apparent density across pipe.)

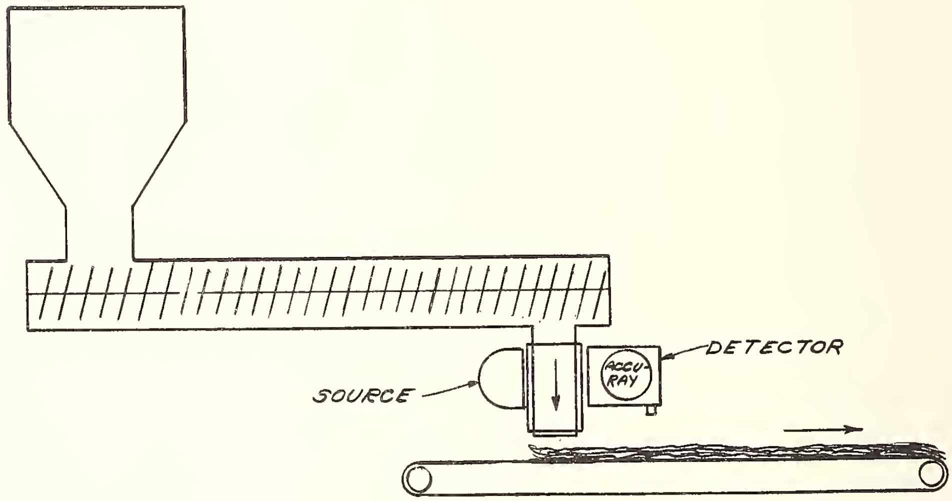


Fig. 34 Nuclear Density Gage Measuring Mass Flow of Solids in Free Fall

V. LIQUID LEVEL

Primary Detectors

The primary detectors are pressure gauges and the diaphragm box.

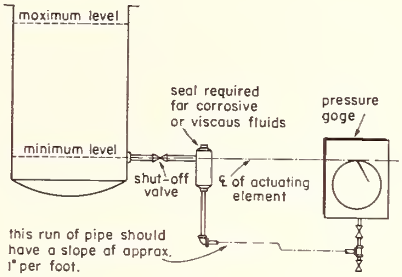


Fig. 35 Liquid Level Measurement by Pressure Gage Method

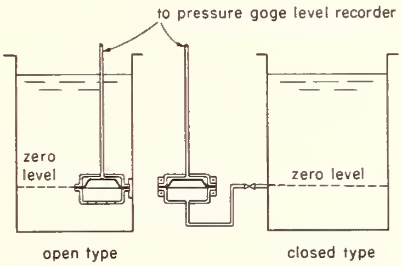


Fig. 36 Liquid Level Measurement by Diaphragm Method

Static Pressure Method

The static pressure measurement is used in measuring the level in "open" or "closed" vessels. This method is based upon the fact that the static pressure exerted by any liquid is directly proportional to the height of liquid above the point of measurement, irrespective of volume. Thus any instrument which measures pressure can be calibrated in terms of the height of a given liquid and used to measure liquid level in vessels under atmospheric pressure. The water columns will add a quadratic lag term to the transfer function of the detector.

In closed vessels, however, the pressure existing above the liquid is added to the liquid head and must therefore be compensated for.

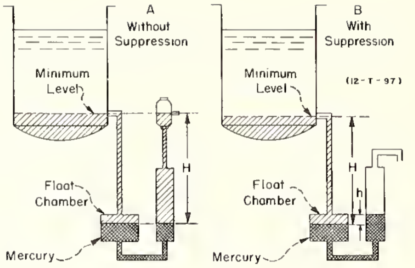


Fig. 37 Liquid Level System with and without Suppression

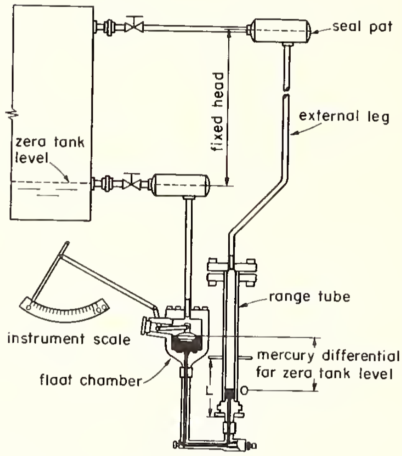


Fig. 38 Liquid Level Measurement in Closed Vessel

The Air Purge Method

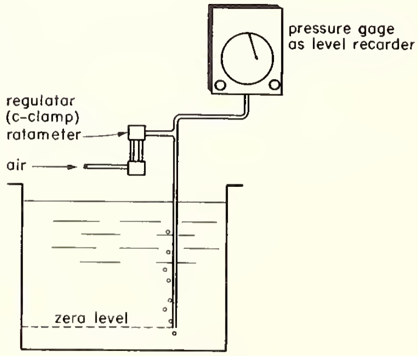


Fig. 39 Liquid Level Measurement by Air Purge Method

An air line is immersed in the liquid to the minimum level as shown with the pressure and volume of the air supply controlled by a regulator so that slow bubbling occurs when the vessel is filled. The pressure in the air line is then equal to the back pressure exerted by the head of the liquid. Measurement of the air pressure is equivalent to measurement of the static pressure of the liquid, i.e. the liquid level. The Dynamic Response is essentially that of the pressure transmitter used.

The Gamma Ray Method

This method is based on the change in the number of gamma rays that penetrate a layer of liquid. As the thickness of the layer increases, the number decreases.

A radiation source unit is used, and its output is measured by a high-pressure, high-efficiency ionization chamber. The ionization current developed is directly proportional to the detected radiation and is amplified for measuring purposes. This is an extremely accurate method and is frequently used on liquids difficult to measure since the source and detector can be located on the outside of the vessel. (See Fig. 34)

Liquid Level by Means of Noise Level Detection

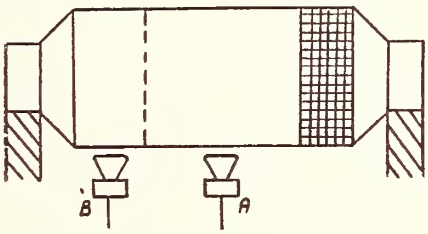


Fig. 40 Liquid Level Detection by Noise Level Measurement

Used frequently on grinding mills to maintain an optimum feed rate. A microphone located near the mill is actuated by the sound, extraneous noises are filtered out, and the transient noise pattern is integrated and converted to a usable current or voltage. The response is essentially instantaneous.

Liquid X Solid Level by Electrical Capacitance

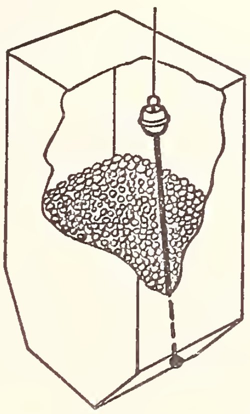


Fig. 41 Liquid or Solid Level Detection by Electrical Capacitance Measurement

The sensing probe assembly consists of a tensioned insulated wire fixed securely at the top and bottom of a container. It operates on the principle of a change in electrical capacitance, the material in the container forming the dielectric between two plates of a capacitor as represented by the probe and the container wall. The response is essentially instantaneous.

VI. SPEED MEASUREMENT

Electric Generator Tachometers

The most commonly used primary elements to measure speed are the AC or DC Tachometer generators. Coupled to a shaft, each will develop an emf which is proportional to shaft speed.

Proximity Probe Method

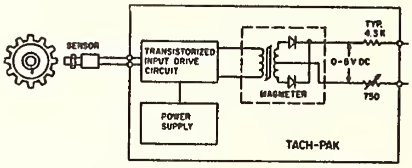


Fig. 42 Speed Measurement Using Proximity Probe and Pulse-Current Convertor

A magnetic proximity pickup senses the rotation of a gear and produces a small signal. This signal, alternating at a frequency dependent on the number of teeth on the gear and its speed of rotation is converted by a frequency discriminator to a proportional DC output.

VII. DENSITY OF LIQUIDS AND SLURRIES IN PIPES

The gamma ray method, used most frequently has been previously described.

VIII. DEWPOINT DETECTING ELEMENTS

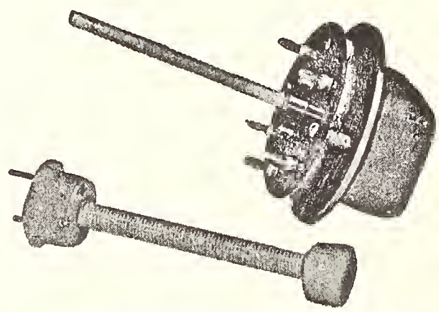


Fig. 43 Dewpoint Detecting Elements

The Dew Probe consists of bifilar wire electrodes wound on a cloth sleeve, which covers a hollow tube or bobbin. The sleeve is impregnated with a lithium chloride solution which is allowed to dry. The bifilar electrodes are not interconnected but depend on the conductivity of the atmospherically moistened lithium chloride for current flow.

At values of 11% RH and below, the lithium chloride becomes conductive and an electric current passes through and heats it. Some of the moisture is thereby evaporated until an equilibrium temperature is reached on the bobbin. At this temperature the R.H. of the air next to the heated bobbin is 11%. The equilibrium or bobbin temperature is thus related to the dewpoint temperature of the air. A filled thermometer bulb or a resistance thermometer is mounted inside the bobbin to measure the cavity temperature.

IX. MEASURING ELEMENTS AND TRANSMITTERS

Having studied a few of the primary detecting elements we are now ready to discuss briefly the next box on our basic control system, which is the measuring element. This is the portion of the feedback elements, which converts the signal from the primary detecting element to a form comparable with the reference input, usually the set point located in the indicating or recording device.

It is quite difficult to say, where we will find the measuring element. In some cases it is incorporated into the indicating and recording device, in others we will find, that due to the remote location of the comparator the measuring element is built into the same housing with the primary detector.

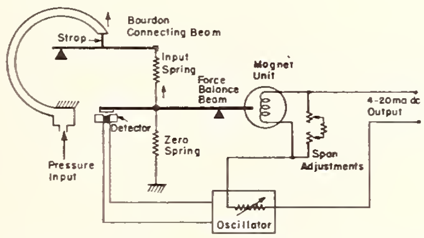


Fig. 44 Method of Operation, PP/1 transmitter

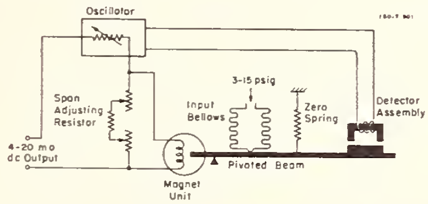


Fig. 45 Method of Operation P/1 Transducer

See Fig. 46 for response of P/1 transducer. Many times it becomes necessary to change the signal from one form to another, such as pneumatic 3-15 psi to electric 4-20 MA in order to make it suitable as an input to an electronic comparator. Transducers and converters have been designed for this purpose.

There are a number of devices on the market which will permit you to pneumatically divide, multiply, ratio, add and subtract. The response for ratioing is given by Fig. 47.

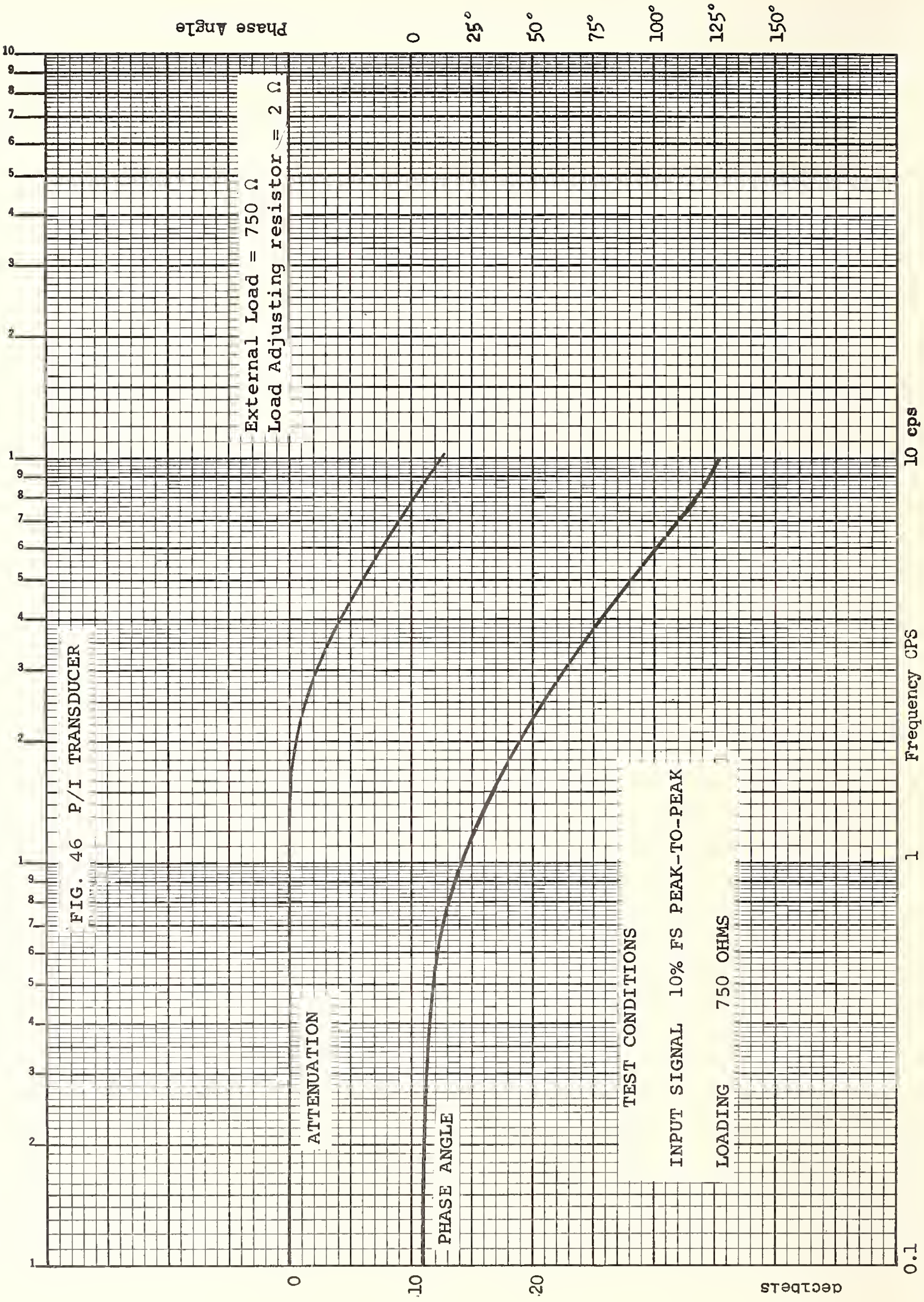


FIG. 46 P/I TRANSDUCER

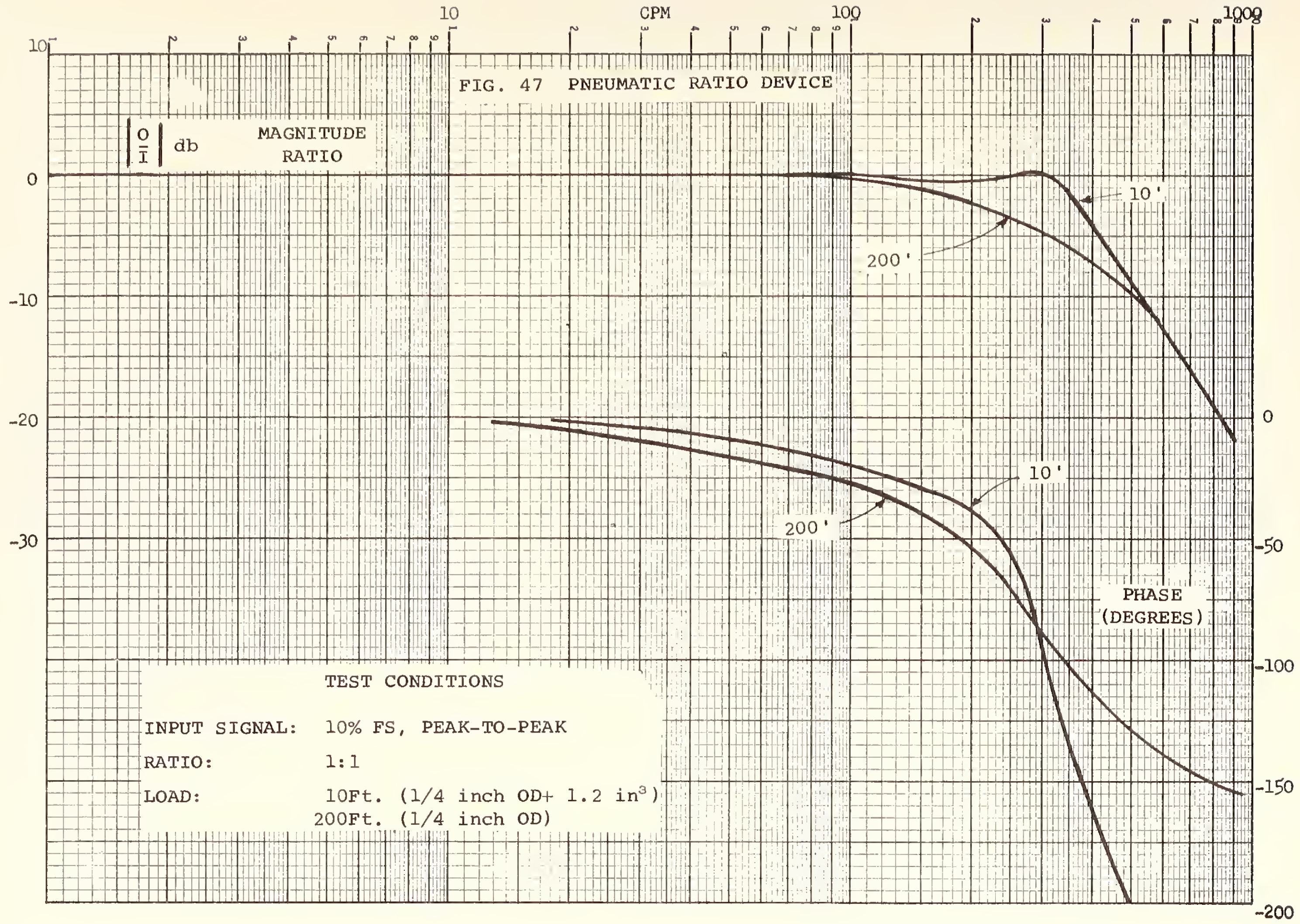
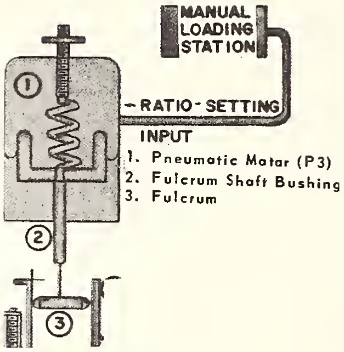
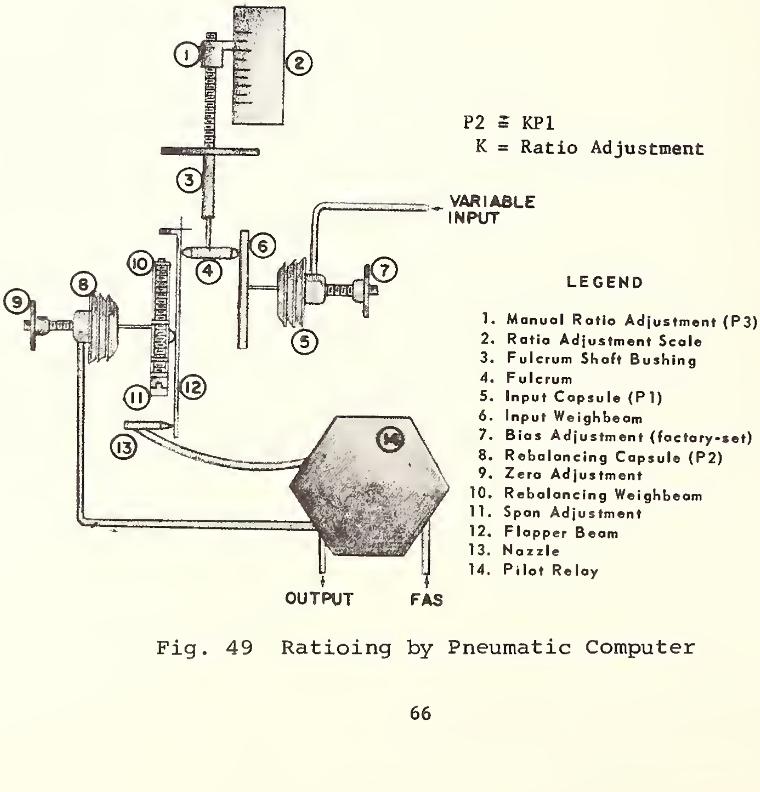


FIG. 47 PNEUMATIC RATIO DEVICE



P2 = KplP3
K = Constant of Mechanical Prop.

Fig. 48 Multiplication by Pneumatic Means



P2 ≅ KP1
K = Ratio Adjustment

Fig. 49 Ratioing by Pneumatic Computer

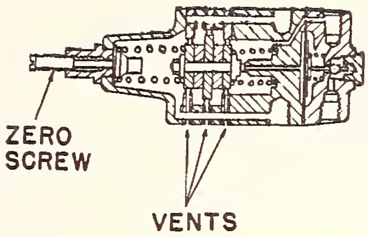


Fig. 50 Adding and Subtracting Relay

There are many electronic solid state auxiliary devices which allow you to:

Add

Subtract

Compare voltages to detect alarm and limit points

Maintain system control at the last operating point upon loss of input

Extract square root - (Fig. 51)

Ratio - (Fig. 52)

Integrate - (Fig. 53)

Compensate for residence time in grinding loop - (Fig. 54)

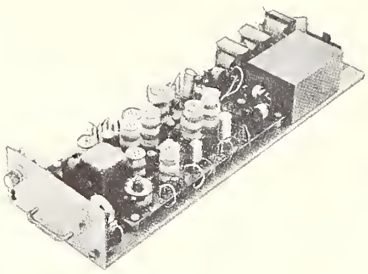
Make up feed losses

Multiply

Automatically select control

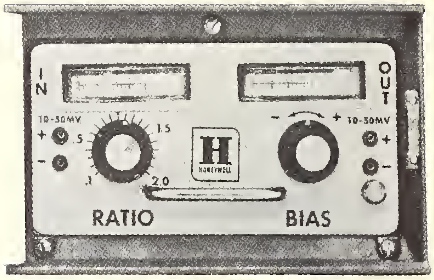
Delay signal - (Fig. 55)

Override set point signal



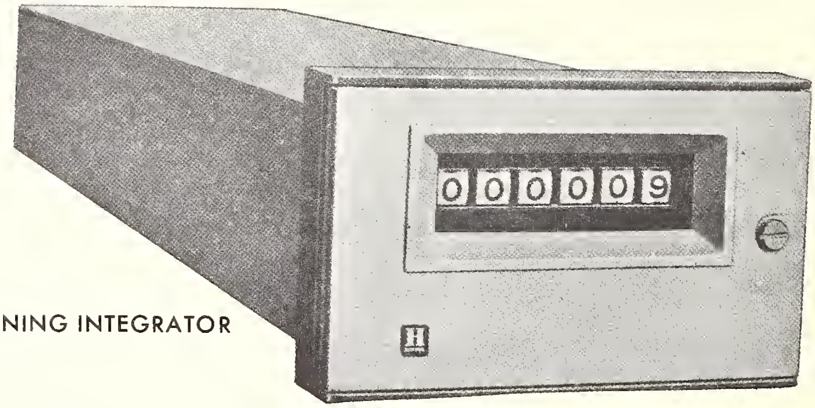
SQUARE ROOT EXTRACTOR

Fig. 51



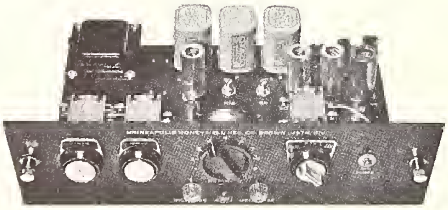
RATIO/CASCADE AUXILIARY

Fig. 52



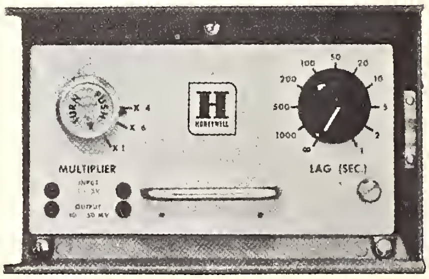
SCANNING INTEGRATOR

Fig. 53



DEVIATION PROPORTIONAL
PULSE CONTROL UNIT

Fig. 54



SIGNAL DELAY RELAY

Fig. 55

X. CONTROLLERS

The basic characteristic of an automatic controller is the manner in which it acts to restore the controlled variable to the desired value. This is called the mode of control. The common modes of control are:

1. Two Position
2. Floating or Single Speed Floating Control
3. Proportional
4. Proportional + Reset
5. Proportional + Reset + Rate

Two Position Control

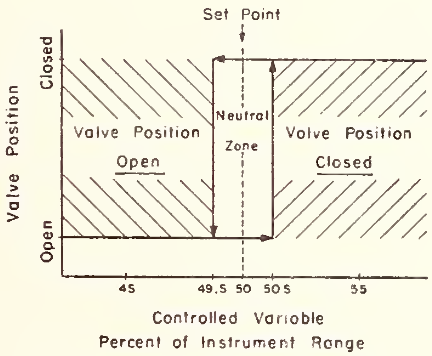


Fig. 56 Two Position Control

In two position control, whenever the controlled variable deviates a predetermined amount from the set point, the controller moves the final control element to either of two extreme positions.

Because of its simplicity, two position control is very popular and often adequate for process regulation. In general, it functions satisfactorily if the process has a slow reaction rate, minimum transfer lag or dead time, and the two extreme positions are adjusted to permit an input just slightly above and below requirements for normal operation.

Floating or Single Speed Floating Control

Floating or single speed floating control moves the final control element at a constant speed, in either direction, whenever the controlled variable deviates a predetermined amount from the set point. The floating controller changes its speed to on and off, while the two position controller changes the position of the final control element from on to off.

The outstanding advantage of floating control over two-position is that gradual load changes can be counteracted by gradual shifting of the valve position, thus minimizing cycling.

Floating control cannot be used where there is any significant lag, or where load changes even though small, occur rapidly.

Proportional Control

In proportional position, mostly called proportional or single-mode control, there is a fixed linear relationship between the value of the controlled variable and the position of the final control element. The proportional controller moves the final control element to a definite position for each value of the controlled variable.

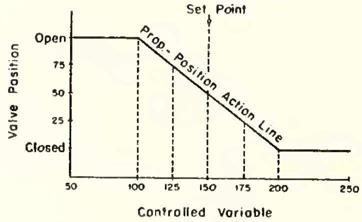


Fig. 57 Proportional-position Control

The proportional controller provides smoother control than the more simple modes because it can move the control valve to intermediate positions that are proportional to the deviation from set point.

The proportional controller, however, can move the control valve to one, and only one, position for any given value of the controlled variable, and therefore is only applicable where the load does not change. It is unable to change valve position for a change of load.

This characteristic of proportional-control results in offset, which is a sustained deviation of the controlled variable from the set point as the result of a sustained process load change. Most proportional controllers have a manual reset adjustment that eliminates offset by shifting the proportional band about the set point.

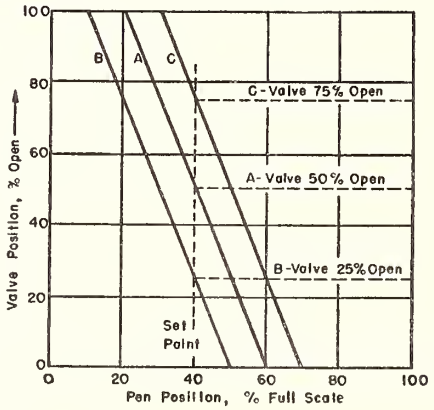


Fig. 58 Effect of Manual Reset

Proportional control is used where process capacitance is relatively large, process reaction rate relatively slow and process lag and dead time relatively small. All of these characteristics promote stability and permit the use of a narrow proportional band, which leads to faster corrective action and less cycling. Under these conditions proportional control can even overcome small load changes.

The wider the proportional band, the greater the offset when there is a load change.

Proportional + Reset Control

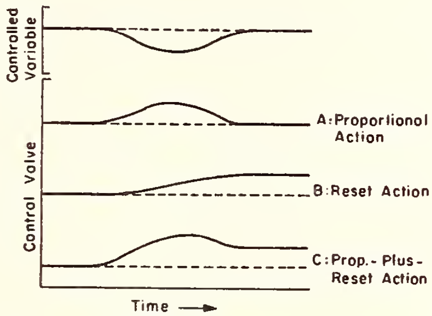


Fig. 59 Proportional + Reset Control

In the proportional + reset mode of control, reset is automatic. As soon as the controlled variable deviates above and below the set point, there is a gradual and automatic shift of the proportional band to bring the variable back to the set point causing changes in valve position to accommodate for load changes.

The illustration shows the changes of a control valve position during a sustained change in process load under:

A - Proportional Action

B - Reset Action

C - Proportional + Reset Action

The primary advantage of proportional-plus-reset control action is the reduction of offset. It can be used even when process capacitance is small, reaction rate if fast and load changes are large.

Rate

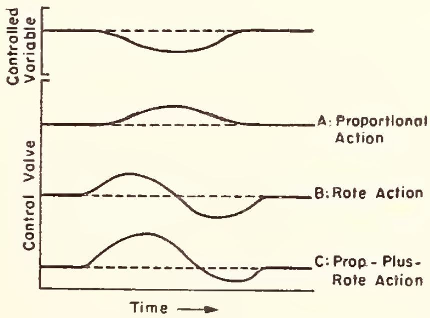


Fig. 60 Proportional-plus-Rate Control

Rate action provides a continuous relationship between the rate of change of the controlled variable and the position of the final control element. Rate action never exists alone, but in combination with proportional or proportional + reset actions. It is a temporary overcorrection proportional to the amount of deviation.

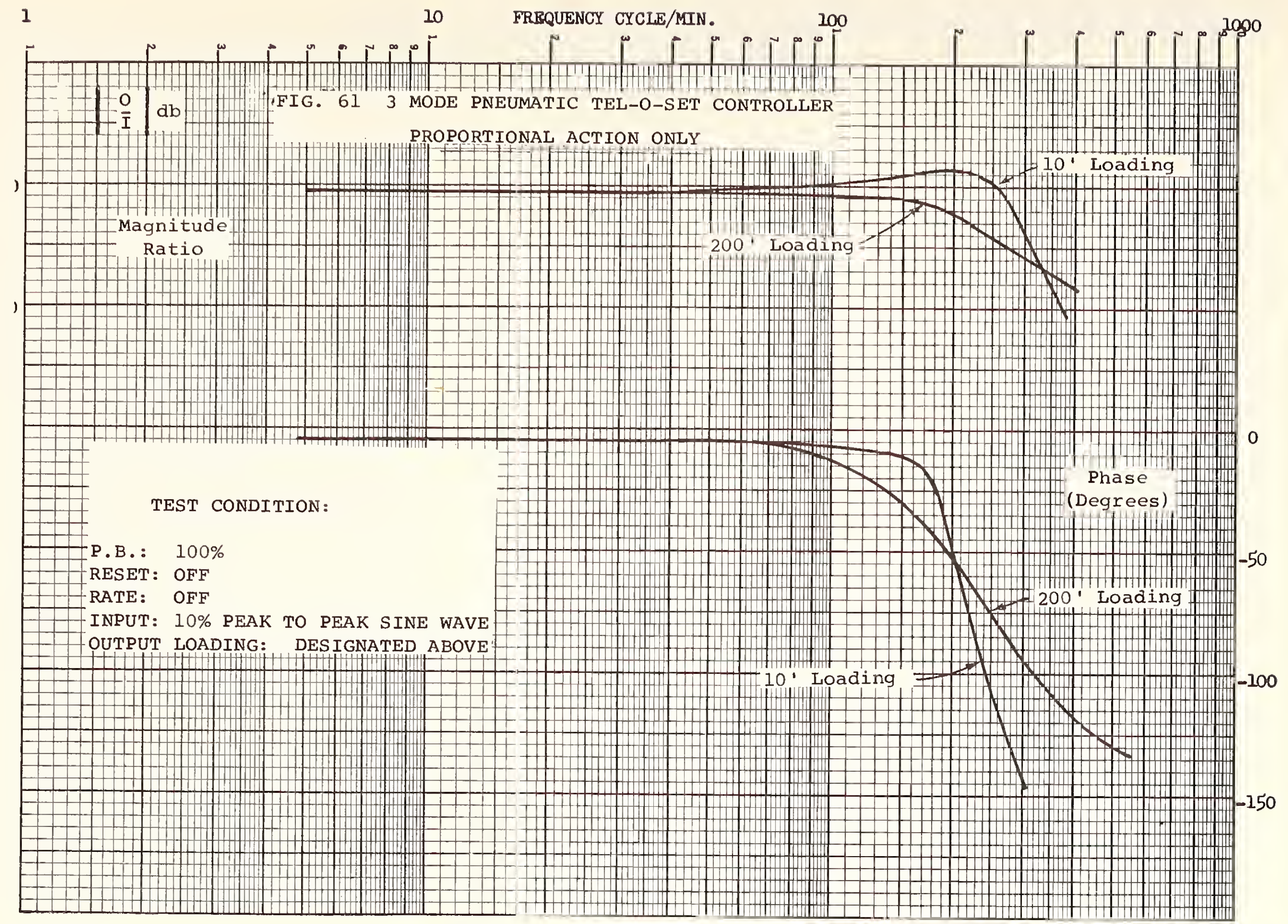
The illustration shows how a proportional + rate controller changes the position of the control valve when the controlled variable deviates from set point.

The diagram breaks the control action into its two components: There is no change in the proportional response A.

Rate response B causes an overcorrection first in one direction then in the other.

C shows that the return of the controlled variable to the set point is sooner than it would without rate action.

Rate action is introduced on processes with large dead time, 2 minutes or more, or large transfer lag that are difficult to control with the proportional + reset mode. Especially on those, which are controlled with a wide proportional band and very slow reset rate. With load change, there is a rather wide deviation and it takes a long time for the controlled variable to return it to the set point. The addition of rate will often solve this problem. Rate time is the time interval by which rate action advances the effect of proportional action upon the final control element, Fig. 61, 62, and 63.



FREQUENCY CYCLE/MIN.
1
10
100
1000
O/I db
FIG. 61 3 MODE PNEUMATIC TEL-O-SET CONTROLLER
PROPORTIONAL ACTION ONLY
Magnitude Ratio
10' Loading
200' Loading
TEST CONDITION:
P.B.: 100%
RESET: OFF
RATE: OFF
INPUT: 10% PEAK TO PEAK SINE WAVE
OUTPUT LOADING: DESIGNATED ABOVE
Phase (Degrees)
0
-50
-100
-150
200' Loading
10' Loading

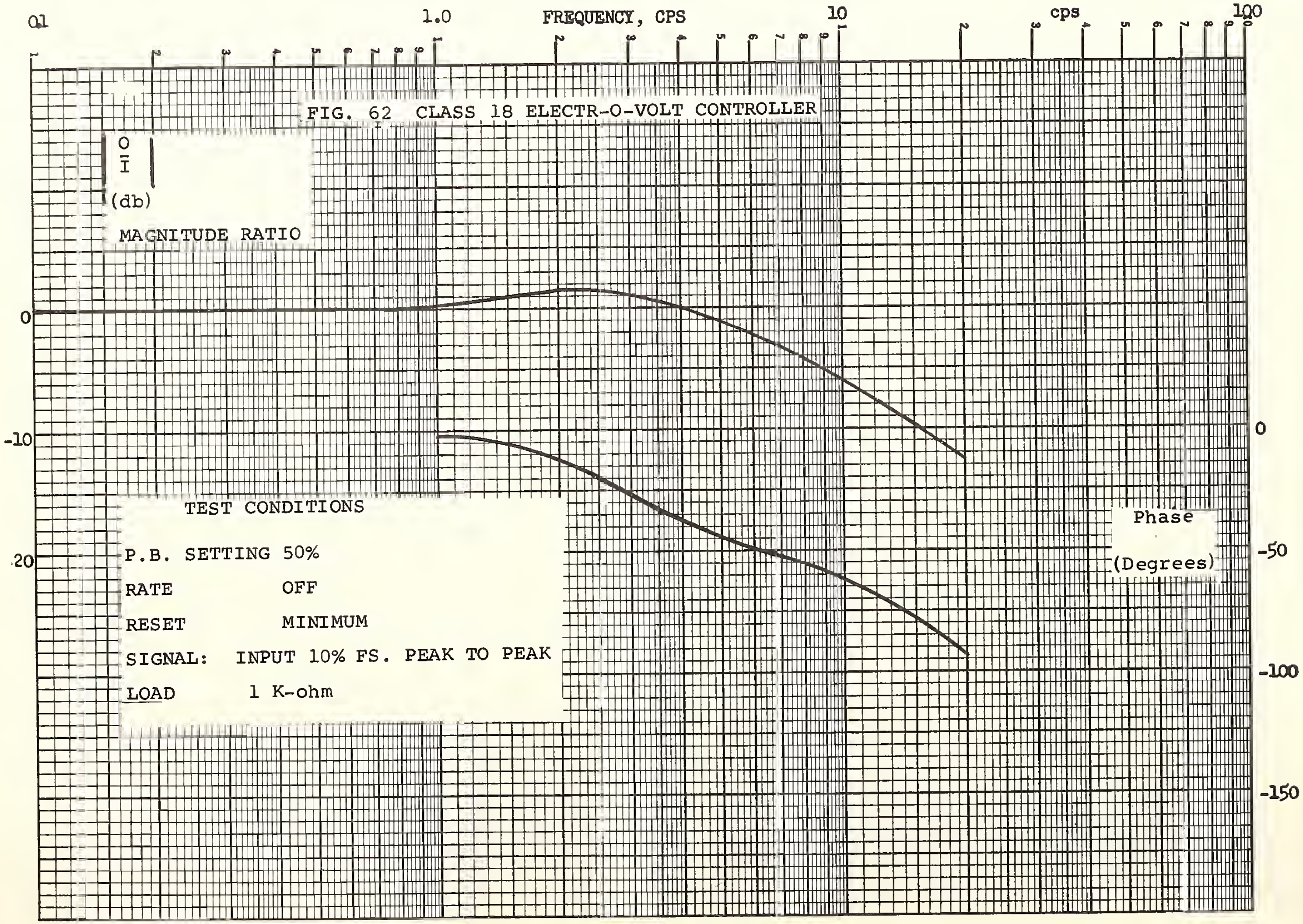


FIG. 62 CLASS 18 ELECTR-O-VOLT CONTROLLER
FREQUENCY, CPS
cps
0.1
1.0
10
100
$\left|\frac{O}{I}\right|$ (db)
MAGNITUDE RATIO
0
-10
-20
TEST CONDITIONS
P.B. SETTING 50%
RATE OFF
RESET MINIMUM
SIGNAL: INPUT 10% FS. PEAK TO PEAK
LOAD 1 K-ohm
Phase (Degrees)
0
-50
-100
-150

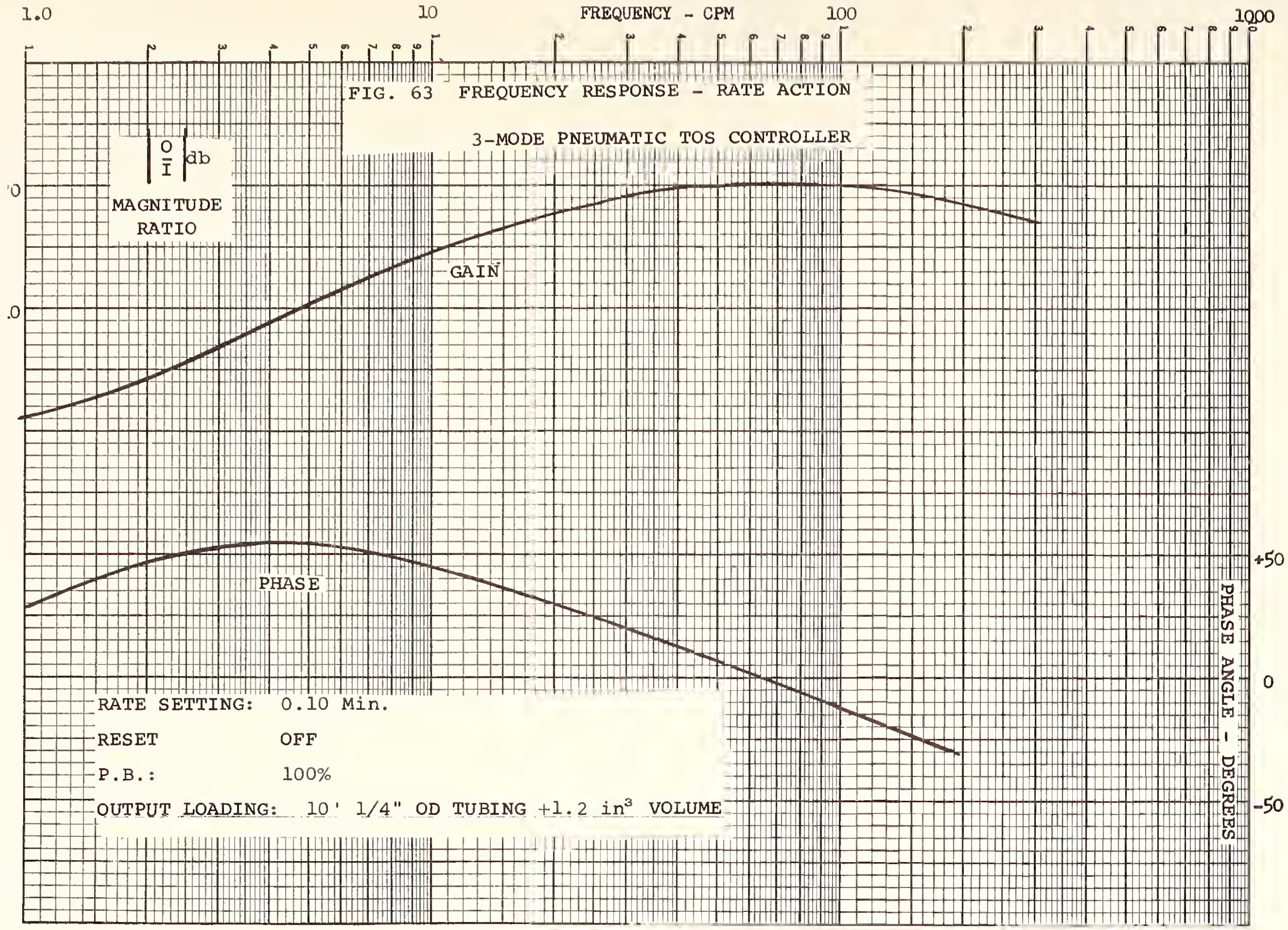


FIG. 63 FREQUENCY RESPONSE - RATE ACTION

3-MODE PNEUMATIC TOS CONTROLLER

XI. COMPLETE CONTROL LOOPS

It is not the scope of this paper to discuss the last box of the basic diagram, the final control element. However, examples of two control loops employing some of the components discussed are given.

Example 1:

It is desired on certain coolers to control the Fan Damper from secondary air temperature, and cooler grate speed from Undergraduate pressure. Should secondary air temperature show a consistent deviation from setpoint, the setpoint of the grate speed controller should be automatically changed to correct this condition. This action should not be erratic.

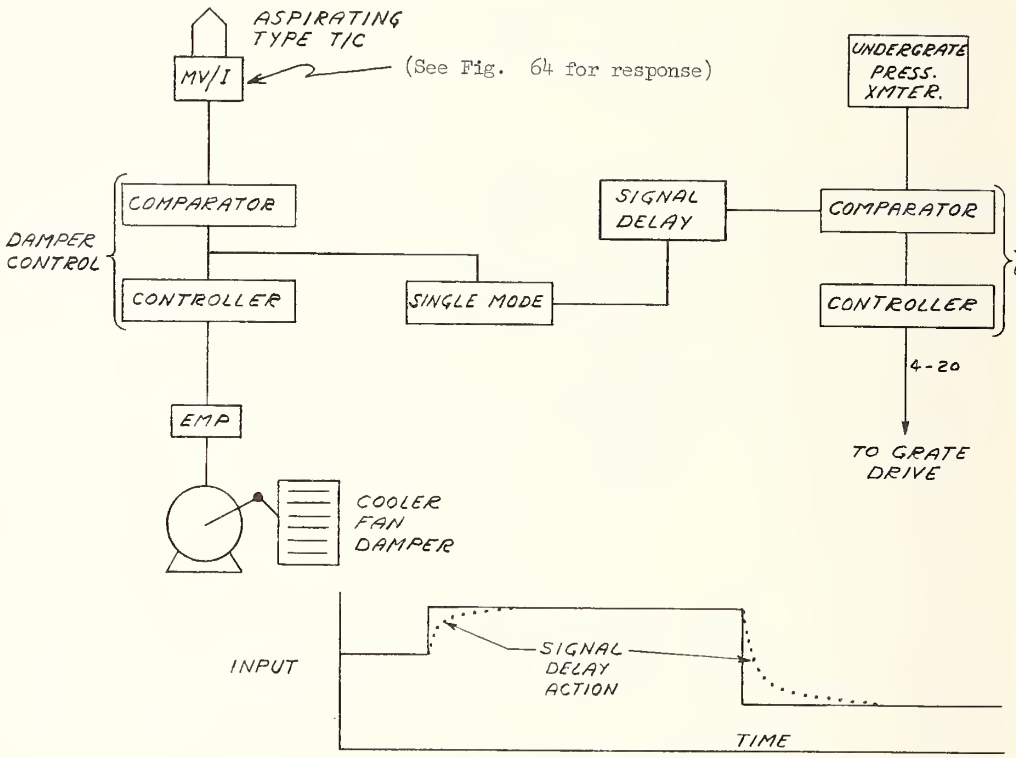


Fig. 65 Typical Closed Loop Control System

Characteristics of this system:

1. Since an ordinary thermocouple in the throat of the cooler would be affected by heat radiation from the clinker, an aspirating type thermocouple is used.

2. If the error signal caused by the deviation of the temperature measured and the setpoint in the comparator are allowed to affect the setpoint of the speed controller directly, the control action would be unstable.

Once the controllable limit of the Fan Damper has been reached, the system would tend to cycle around this limit. Therefore, a time delay unit has been employed, the action of which is very much like an RC network, forward and backward.

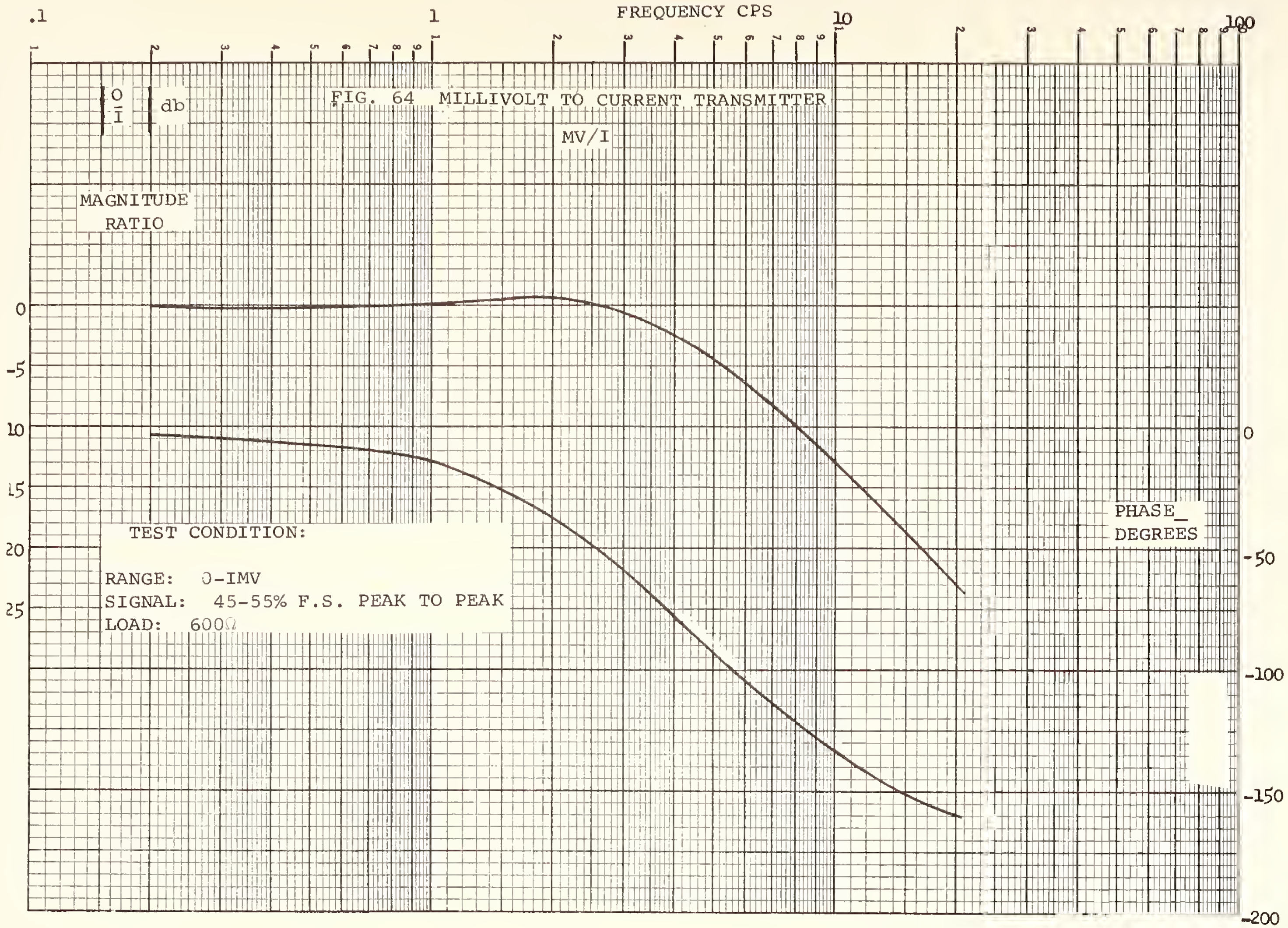


FIG. 64 MILLIVOLT TO CURRENT TRANSMITTER

The single-mode controller is the biasing means of the electrical setpoint and acts otherwise like a controller with proportional band.

Example 2:

A system to blend 3 materials is required. The material should be fed to a 3-compartment ball mill in a closed circuit. Thru-put measured by a sonic device should control feed and primary water additions. Slurry should be held at 35% water by weight ahead of a DSM screen.

System should be of highest accuracy possible.

Flow Diagram of System:

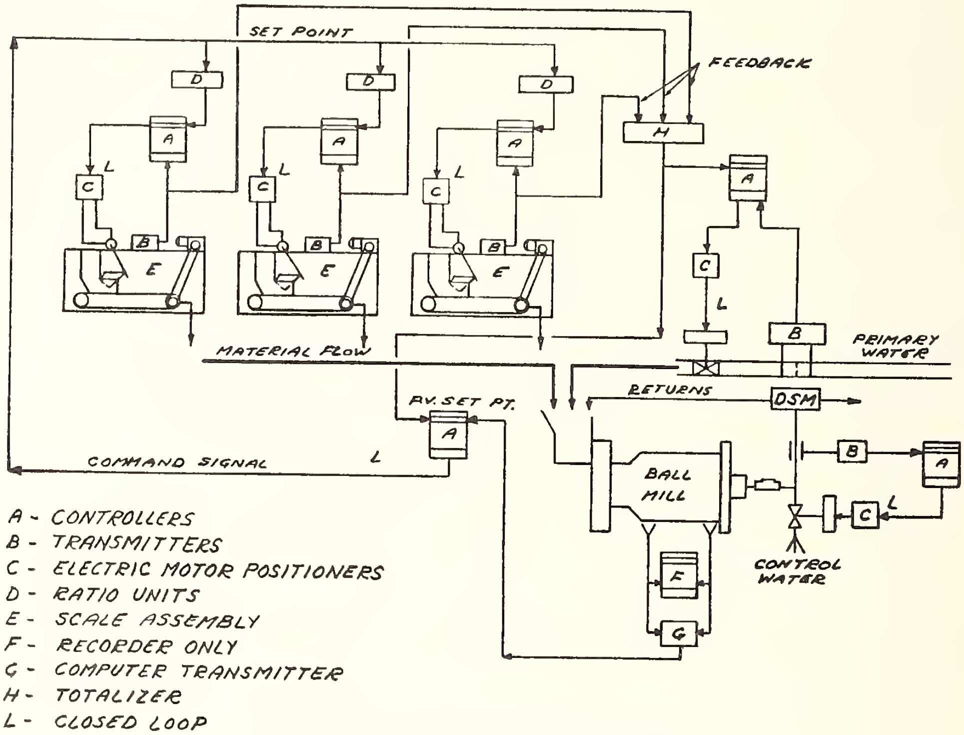


Fig. 66 Typical Blending System

Basic characteristics of system:

1. Weight inaccuracy due to inherent moisture in the materials.
2. Variation in the material size.
3. A lag produced by the material residence time in the grinding mill.
4. Variation in circulating load returns.

The worst condition would exist if the materials to be blended have a different inherent moisture content. The solution for this problem would be to preheat the materials or take the average of separate moisture measurements. Neither solution is satisfactory in this process because of the high cost of moisture control. Frequently it is decided to not control moisture and it demands an increase in the accuracy of the weighing equipment which furnishes the feedback signal to the control loop. If this signal is of poor quality, the entire system becomes affected by it. Variation in material sizes results in considering only the weighing method. Residence time in the Ball Mill results in considering the circulating load returns, and adds an unknown amount to the primary water. This will have to be considered if it is necessary to control the required percentage of water by weight of the slurry.

Bibliography

1. D. P. Eckman: Automatic Process Control (John Wiley & Sons)

2. J. G. Truxal: Control Engineer's Handbook (McGraw-Hill)

3. Fundamentals of Industrial Instrumentation (Honeywell)

4. G. K. Tucker & D. M. Wills: A Simplified Technique of Control System Engineering (Honeywell)

5. T. C. Linahan: The Dynamic Response of Industrial Thermometers in Wells (ASME Trans. 1956 Vol. 78, No. 4)

6. D. P. Eckman: Industrial Instrumentation (John Wiley & Sons)

7. D. M. Considine: Process Instruments & Controls Handbook (McGraw-Hill)

8. Orton St'd Pyrometric Cones: The Properties and Uses of Pyrometric Cones

Application of Computers in Design and Control of Systems

Joseph L. Hammond, Jr.

1. Introduction

The inherent complexity of most problems which are of interest to ceramic engineers makes tractable mathematical models almost impossible, and thus analytical results have historically been of considerably less interest than practical experience. The rapidly advancing fields of automatic control and automated design, however, depend to a large extent on the use of mathematical models for the problems of interest. Perhaps the key to making automatic control and automated design possible in the ceramic industry can be the modern computer which has the ability to cope with mathematical models far more complex than can be handled efficiently with hand calculation.

To use a computer, some sort of mathematical formulation of the problem must be available. However, the present state of the art is such that many formulations such as nonlinear equations, large sets of simultaneous equations, equations with random parameters etc., can be handled numerically with a computer even though such formulations are not analytically tractable. The problem of choosing a mathematical model is frequently difficult and of course has important significance on the overall results. In order to approach the use of computers intelligently, the engineer must be cognizant of the properties of the chosen computer when he formulates the system equations. Only in this way can full use be made of the capabilities of the machine. In later sections of this paper this point will be illustrated by specific examples.

The role of the computer in systems work has two distinct parts: (1) the computer can be used as an element in a control system, or (2) it can be used as a tool in system design. The use of computers in system design can be further divided into several catagories, namely (a) their use in analyzing systems and system components, (b) their use in simulating systems, and (c) their use in automated design procedures.

In the discussion below each distinct use of the computer will be given separate attention. Before discussing the uses of computers, however, the two computer types - analog and digital - will be described, and the essential features of methods for solving ordinary differential equations will be reviewed as an illustration of how computers are programmed.

2. Characteristics of Analog Computers

. An analog computer is distinguished by the fact that every variable in a problem to be solved is represented in the computer by the magnitude of some physical quantity. Electrical voltage is currently the most popular quantity and this will be assumed here for definiteness. In almost all cases the independent problem variable is represented by time in the computer. Thus the result of a computation is one or more voltages which vary with time in a manner approximating the solution to the given problem. These voltages can be recorded to give a permanent graph of the solution to the problem.

Since the dependent problem variables are represented by voltages, the analog computer must have available elements to operate on these voltages in the manner dictated by the equation. Thus analog elements perform the following operations:

1. Integration
2. Addition (subtraction)
3. Multiplication of a variable by a constant
4. Multiplication of two variables together
5. Generation of appropriate functions of dependent variables.

Note that differentiation is not normally included since it is a noise accentuating operation. A true analog computer with no digital accessories is handicapped by being able to perform only a limited number of logical operations such as "when voltage A exceeds voltage B then do operation C, otherwise do operation D."

The accuracy of an analog computer is inherently limited by the fact that problem variables are represented by physical quantities. Experimentalists in all areas realize that measurement accuracies cannot be made arbitrarily high, and that devices responding to physical variables are limited in operating range. Good general purpose analog computing equipment will have a typical component accuracy of from 0.01 to 0.1% of full scale, and a linear range on the order of 1000 to 1. This limitation on accuracy makes an analog computer best suited to solving ordinary differential equations in problems where "engineering" accuracy is adequate.

As stated above, the independent variable in an analog computer is time. Fortunately, however, actual time in the analog computer can be proportional to the independent problem variable and does not necessarily have to be equal to it. The same comment can be made relative to the voltages representing dependent problem variables. The constants of proportionality for the independent and dependent variables are called scale factors, and their use extends greatly the useful range of the computer. Applying the scale factors to the problem equation is essentially a change of variables which results in a "machine equation" for use in the computer.

The basic steps in using an analog computer are shown in Figure 1. First scale factors are chosen and applied to the problem equation. The result of this change of variables is a machine equation which is solved on the computer with voltages representing the dependent machine variables and time representing the independent machine variable. The computer outputs are plotted vs time on an appropriate recorder and finally these plots are calibrated in problem variables by use of an inverse scale factoring procedure.

3. Characteristics of Digital Computers

As contrasted to analog computers, the dependent variables at any time in a digital computer are represented as binary numbers by a collection of computer elements which have only two stable states. The nature of the digital computer is such that the variables cannot change continuously, and thus differential equations, for example, must be approximated by difference equations for which the independent variables progress in steps of a nonzero length.

A digital computer performs five basic operations, (1) input from tape, cards, etc., (2) arithmetic operations on data in the machine, (3) storage, (4) output to typewriters, etc., and (5) logical control of the sequence of operations which are performed. The actual computation is done by proper sequencing of arithmetical operations on the input and stored data. The basic arithmetic operation is addition, but this can be combined with storage to perform multiplication, division, integration and other operations. A block diagram showing these operations is given in Figure 2.

The versatility of the modern digital computer is a result of the speed with which the basic operations of addition, storage, and shifting of information can be accomplished. For present day machines the time required for a single operation is on the order of 10^{-6} sec. To make proper use of this type of computer, however, involved programming is required to convert the required calculations into the proper sequences of the basic machine operations. Fortunately much of the detailed instruction can be preprogrammed once and for all into the machine memory or stored on paper tape. A number of so called interpretive programs such as Fortan are available which preprogram the machine to receive directly instructions to perform operations such as multiplication integration etc., which are ultimately carried out as sequence of the simpler operations.

As pointed out above the digital computer works with numbers, usually in the form of binary digits. Thus the dependent variables at a particular multiple of the step size of the independent variable are represented by binary numbers at appropriate locations in the machine. The high accuracy of the digital computer results from the fact that each digit of the binary number is represented by one of two possible states. For example, 1 could be represented by +100 volts and 0 by -100 volts. An error due to noise is seldom so large that +100 volts is interepreted as -100. Thus the binary numbers can be processed almost without error and the primary source of error is rounding off to a fixed number of binary digits. Even here the digital machine has the flexibility of being able to swap computing time for accuracy by using a large number of binary digits. The accuracy of simple operations such as addition, multiplication etc., can thus be made almost arbitrarily high. However, as in the case of the analog computer, it is a difficult matter to determine overall accuracy for a non trivial problem in terms of the accuracy of the elementary operations, and of course, there are many problems which cannot be solved accurately on a digital computer.

In addition to high accuracy the digital computer has great flexibility in performing logical instructions. This fact makes automated design feasible since, for example, a digital computer has the capability of solving a set of equations for a system, comparing the result to a performance criterion, and then changing the system parameters and repeating the calculation.

4. Computer Solution of Differential Equations

Many systems of interest to ceramic engineers can be described by ordinary differential equations. Those systems which cannot be exactly described in this way can frequently be approximated by ordinary differential equations. Thus this very important class of equations is chosen to illustrate the technique of machine computation.

Most systems of ordinary differential equations which can be solved for the highest derivatives can be reduced to a system of the form

$$\begin{aligned} dx_1/dt &= f_1[t, x_1(t), ----, x_n(t), y_1(t), ---, y_m(t)] \\ &\vdots \\ dx_n/dt &= f_n[t, x_1(t), ----, x_n(t), y_1(t), ---, y_m(t)] \end{aligned} \tag{1}$$

with appropriate initial conditions. In Equation (1) the f_i are real valued functions of the indicated arguments, the $y_i(t)$ are known forcing functions, and n is the order of the given system of equations. For a physical system the variables x_1, --- , x_n are just the state variables discussed in an earlier paper, and Equation (1) can be thought of as the description of a physical system in state variable form.

To illustrate the procedure for obtaining a numerical solution to a set of equations such as Equation (1), consider the case of only one state variable. For this case the equation is simply

$$dx/dt = f[\, t, x(t), y(t) \,] \tag{2}$$

with an initial condition given by

$$x(t = t_o) = A.$$

A numerical solution to equation (2) is obtained when x is specified for values of t over the range of interest. It is instructive to note that the given equation prescribes the derivative or slope of the desired curve at every value of t. Thus in effect the differential equation prescribes the locus of tangents to the desired curve. This fact can be used in obtaining a numerical solution to the equation. To begin with, x is known to be A at $t = t_o$. At this point Equation (2) gives the slope of the curve as

$$\frac{dx}{dt}\Big|_{t = t_o} = f[\, t_o, A, y(t_o)\,]. \tag{3}$$

The problem is then to obtain the values of x at values of t other than t_o. A convenient approximation is to assume that the actual x(t) vs t curve coincides with the tangent to the curve at $t = t_o$ for t values near zero. This results in the expression

$$x^*(t) = \int_{t_o}^{t} f[\, t_o, A, y(t_o)\,]\, dt + A \tag{4}$$

which is a good approximation to x(t) for t "close" to t_o, say in the range $t_o \le t \le t_o + \varepsilon$. Now the approximation

$$x^*(t) = \int_{t_o+\varepsilon}^{t} f[\, t_o+ \varepsilon, x^*(t_o+ \varepsilon), y(t_o+ \varepsilon)\,]dt + x^*(t_o+ \varepsilon) \tag{5}$$

can be used for t in the range $t_o + \varepsilon \le t \le t_o + 2\varepsilon$, and the process can be continued to determine x*(t) over the whole range of interest.

Approximations of the general type of x*(t) are used in both analog and digital computation. Questions of convergence of x*(t) to x(t) of course arise, and such considerations lead to more involved approximations but the basic idea is preserved.

The circuit of Figure 3 is used to generate x*(t) with an analog computer. The circuit operates as follows. At $t = t_o$ the integrator output A, a voltage corresponding to t_o, and a voltage corresponding to $y(t_o)$ are the inputs to a collection of elements which produce $f[t_o, A, y(t_o)]$. As time progresses, the inputs to this "function generator" change so that f[t,x*(t),y(t)] is produced and integrated to produce x*(t). The "closed loop" feature of this circuit, namely the fact that the integrator output, x*(t), is one input to the function generator which produces the integrator input to correct x*(t), is characteristic of any machine solution of a differential equation.

In a digital computer the computation of x*(t) is accomplished by essentially the same logic as shown in Figure 3. For this computer, however, the integral in Equations (4) and (5) is also approximated, and the independent variable is made to progress in discrete steps, say of value ε. Thus Equation (4) becomes

$$x^*(t_o + \varepsilon) = \varepsilon f[t_o, A, y(t_o)] + A, \tag{6}$$

and the general expression for any discrete t value is given by

$$x^*[t_o + (i+1)\varepsilon] = \varepsilon f[t_o + i\varepsilon\, , x^*(t_o + i\varepsilon), y(t_o + i\varepsilon)] + x^*(t_o + i\varepsilon) \tag{7}$$

Equation (7) is a difference equation which can be solved by the computer circuit of Figure 4. In this circuit starting at $i = 0$, the values of $x^*(t_o + i\varepsilon)$, $y(t_o + i\varepsilon)$, and $t_o + i\varepsilon$ are stored in the memory and used at each step to compute first $f[t_o + i\varepsilon, x^*(t_o + i\varepsilon), y(t_o + i\varepsilon)]$ and then the corrected approximation $x^*[t_o+(i + 1)\varepsilon]$. The quantity $x^*(t_o+ i\varepsilon)$ can be printed out as the machine progresses through the calculation.

The fact that x^* is a continuous function of t for the analog computer, and a discrete function of t for the digital computer is misleading in comparing the two machines. The analog computer in actuality has an inherent smoothing or averaging effect which causes the continuous output. It is well known, however, that due to limited bandwidth the continuous analog signal contains no more, and usually less, information than the discrete digital output.

The general $n^{\underline{th}}$ order ordinary differential equation, expressed in state variable form by Equation (1), can be solved in the same way as the single variable case. For the n variable case, n circuits of the type shown in Figure 3 or 4 must operate simultaneously. Furthermore each of these circuits is coupled together through the fact that in the general case the value of each function f_i depends not on just $x_1^*(t)$, but on all n state variables as well as additional inputs. Of course convergence problems are more severe in the general case so that practical circuits frequently contain modifications of the basic idea.

An interesting logical simplification of the procedure for the general case results from writing Equation (1) in vector form as

$$d\bar{x}/dt = f[t, \bar{x}(t), \bar{y}(t)] \tag{8}$$

where

$$\bar{x} = \begin{bmatrix} x_1(t) \\ \cdot \\ \cdot \\ x_n(t) \end{bmatrix} ; \quad \bar{y}(t) = \begin{bmatrix} y_1(t) \\ \cdot \\ \cdot \\ \cdot \\ y_n(t) \end{bmatrix} \quad \text{and } d\bar{x}/dt = \begin{bmatrix} dx_1/dt \\ \cdot \\ \cdot \\ \cdot \\ dx_n/dt \end{bmatrix}$$

A circuit showing the logic for solving Equation (8) is given in Figure 5. In this figure the double lines denote the flow of vector quantities, and the complete circuit is a short hand representation for n interconnected subloops of the type shown in either Figure 3 or Figure 4. It is to be understood that $\bar{x}(t)$, the output of the circuit of Figure 5, will be n plots representing each $x_i^*(t)$ vs t for the analog computer, and n sets of data tabulating the $x_i^*(t_o+ i\varepsilon)$ at values of $i = 0, 1, 2, ---$ for the digital computer.

5. Operating Characteristics of Computers

In using a computing machine it is important to take into account the obvious fact that best results are obtained on equations which are matched to the capabilities of the computer. Thus use of a computer is itself a system problem for which the physical process, the mathematical model for the physical process, and the computer must all be considered to obtain best results. For example, a mathematical model should be chosen to not only represent the physical process, but to also be compatible with the computer which will be used to give the numerical result. It should be noted that classical mathematical models in almost all fields are "matched" to analytical computation, and may or may not be appropriate to machine computation.

The important problem of choosing a mathematical model which is matched to both the computer and the physical process is unfortunately still an art not a science. However, some of the characteristics of computing machines which have bearing on the choice of mathematical model can be listed and used as a guide in choosing mathematical models.

First, both analog and digital computers have the following common characteristics:

1. Computers can operate only on numerical equations for which all parameter values are specified. While it is possible to generate a family of solutions for different parameter values, a computing machine cannot produce a general solution in the sense of giving an answer as a function of parameter values.

2. The operation of differentiation accentuates noise in both types of computers and should be avoided.

3. Multipoint boundary value problems cannot be solved directly in most cases.

4. Only continuous functions can be produced as solutions to equations.

5. Partial differential equations must be approximated by sets of ordinary differential equations.

Digital computers have the following additional operating characteristics worthy of noting.

1. This type of machine is well suited to a large number of simple operations.

2. Answers are produced for discrete increments in the independent variable, thus finite, as opposed to continuous, state models are appropriate to this type of computation.

3. Memory is limited to a small fraction of the number of operations which can be performed in a short period of time.

4. Tabular data is easily handled by digital computers.

5. Logical decisions of the type "compare A and B, if A is greater than B do operation C, if not do operation D" can be made easily.

6. The method of operation of a digital computer is seldom related to the physical process being analyzed so that computing accuracy has nothing to do with accuracy in the physical process.

7. Complexity in the sense of the total number of operations does not necessarily decrease accuracy. Such complexity can be had at the expense of computing time with little change in the number of components used.

Analog computers are distinguished by the following characteristics:

1. The linear operations of addition, subtraction, multiplication by a constant, and integration can be performed with high accuracy.

2. Multiplication of two dependent variables, and generation of functions of dependent variables are less accurate operations than the linear ones.

3. Logical operations are not easily performed.

4. Continuous state models are appropriate for this type of computer.

5. Complexity, in the sense of the total number of operations, is obtained by a proportional increase in the number of computing elements. Accuracy is almost always a decreasing function of complexity.

6. For many problems there is a direct correspondence between the analog computing elements and elements of the physical process being analyzed. In such cases, the accuracy with which a given mathematical variable can be computed is comparable with the accuracy possible in generating and measuring the corresponding physical variable. Problems of this type are well suited to the analog computer.

6. Use of Computers in Analysis of Systems and System Components

This application is perhaps the oldest and most direct use of computing machines. The type of problems which arise in this area are as many and as varied as the branches and aspects of engineering. Thus subject to the comments made in section 5, the choice of computer method depends on the particular problem, and little more of a general nature can be said. The remainder of this section will be devoted to a discussion of a specific problem as an illustration of typical procedures.

Consider the problem of diffusion of liquids or gases. The concentration, C(x, t) of diffused substance at a point x and time t is usually assumed to satisfy Fick's law which for a one dimensional problem can be written as

$$\frac{\partial C(x, t)}{\partial t} = \frac{\partial}{\partial x}\left[D[C(x, t)]\,\frac{\partial C(x, t)}{\partial x}\right] \tag{9}$$

with appropriate boundary conditions. In this equation the parameter D(C) is the diffusion coefficient.

Equation (9) can be solved analytically if the diffusion coefficient is independent of concentration or if the equation has certain simple boundary values. From a practical point of view, if a given physical process does not have these simple boundary values, D(C) must be assumed to be constant in order to obtain an analytical solution. For a constant D(C), a general solution to Equation (9) can be obtained. The problem then is to choose an appropriate function of the general solution to match the boundary values, and hence solve a specific problem. The final solution for specific boundary conditions is frequently a complicated function of the general solution to Equation (9) and it is sometimes expedient to employ a computer to evaluate this function at specific values of x and t or to expedite other steps in the procedure. Such a use of the computer implements, and makes practical laborious computation.

Subject to certain restrictions, Equation (9) can also be solved numerically on an analog or digital computer for either constant or nonconstant diffusion coefficient. A machine solution of this type of course involves approximations and imposes restrictions which are different from the corresponding ones for an analytic procedure. In the case being considered the mathematical models which can be handled by machine computation can frequently yield better approximations to the physical process. In particular, these are practical problems with nonconstant diffusion coefficients which can be worked on a computer (see reference 40). Although this class of problem is obviously more involved than the simpler constant diffusion coefficient case, use of the computer in such problems extends the analytical methods.

Direct numerical solution of Equation (9) on either an analog or digital computer involves approximating at least one of the derivatives by a finite difference. On the analog computer the partial derivatives with respect to the space variable x are approximated by differences to reduce the partial differential equation to a system of ordinary equations with time as the independent variable. Physically this is equivalent to separating the axis along which diffusion takes place into a number of cells within which the concentration is assumed to be constant. The simplest approximation of this sort, assuming a constant diffusion coefficient, results in the equation

$$\frac{dC_i}{dt} = \frac{D[C_{i+1} - 2C_i + C_{i-1}]}{(\Delta x)^2}, \qquad (10)$$

$$i = 0, 1, 2, ---, N$$

where

$$\Delta x = x_i - x_{i-1}$$

$$C_i = C(x_i, t)$$

N = number of divisions of x-axis

A similar but more involved expression could be obtained in the case of a nonconstant diffusion coefficient. If a large number of subdivisions of the x-axis is required, it is frequently impractical to solve the resulting large number of equations on an analog computer and a digital computer must be used.

For digital computation the derivative of C with respect to t must also be approximated by a finite difference with the effect that the time axis is also subdivided into cells. This results in the set of equations

$$\frac{C_{i,j+1} - C_{i,j}}{\Delta t} = \frac{D[C_{i+1,j} - 2C_{i,j} + C_{i-1,j}]}{(\Delta x)^2} \qquad (11)$$

$$i = 0, 1, 2, ---, N$$

where

$$t = t_j - t_{j-1}$$

$$C_{ij} = C(x_i, t_j)$$

and other symbols have the significance assigned above. Equation (11) represents the simplest type of difference equation equivalent to Equation (9) assuming a constant diffusion coefficient. Many times in practice, stability of the solution becomes a problem and more sophisticated approximations must be used. Some questions of this sort are discussed in the references cited at the end of the paper.

Note that a mathematical model of the form of Equation (11) could be developed for a nonconstant diffusion coefficient. On the other hand for numerical computation using such an equation both the x-axis and the t-axis must be subdivided into cells in which the concentration is constant. Usually the lack of restriction on D(C) more than overrides the necessity to subdivide into cells; and thus if a properly convergent solution can be obtained, the finite number of data points present in the solution is adequate in normal engineering work.

Example 1: As an illustration of the solution of a diffusion problem on a digital computer, consider a one dimensional problem discussed by Crank (Ref. 39, p. 207). The problem involves a plane sheet of material of thickness 2L in the x-dimension, and infinite extent in the y - and z - dimensions. The plane sheet is immersed in a bath of vapour which diffuses into the material in such a way that the equation

$$\frac{\partial C(x,t)}{\partial t} = \frac{\partial}{\partial x}\left[D[C(x,t)]\,\frac{\partial C}{\partial x}(x,t)\right] \qquad (12)$$

is satisfied. The diffusion coefficient varies with concentration as prescribed by

$$D(C) = D_o e^{2.303\, C/C_o} \tag{13}$$

where D_o is a constant, C is the concentration at any x and t, and C_o is the concentration at each surface of the sheet. It will be assumed that the concentration $x = \pm L$ remains at C_o for all t, and that $C(s,t) = 0$ for $-L < x < L$ and $t = 0$. The physical configuration is shown in Figure 6.

To solve this problem, Equation (12) was reduced to a finite difference equation, analogous to Equation 11, by using a method due to Crank. The nondimensional variables

$$C = C/C_o, \quad X = x/L, \quad T = D_o t/L^2, \quad D^* = D/D_o$$

were introduced to make the results as general as possible. The solutions obtained are given in Figure 7 as a plot of C vs X with T as a parameter. The equations used on the computer and some further details of the solution are discussed in Appendix 1.

7. Use of Computers in the Design of Systems

a. Design by Successive Analysis: The most direct system design procedures are based on the successive analysis of trial designs. The procedure is initiated by choosing a particular set of parameter values and a particular configuration of the system elements. The trial design is then analyzed to see if it meets specifications. If it does not, the system is modified, and a new analysis and performance evaluation is made. The procedure is continued until a satisfactory system is obtained.

For some systems the interaction between elements is so complex that it is necessary to deal almost entirely with the overall system equations for a given trial design. In such cases the calculations can be implemented by a computer which is simply used to solve a sequence of analysis problems.

For other systems the interaction between elements is such that system behavior can be usefully expressed in terms of component element behavior. For this type of system a complete computer program would contain a number of element analysis programs as subroutines. In such cases a given element analysis program becomes an entity in itself, and can be used in the analysis of other systems using the same element. Libraries of programs for the analysis of components can be built up and used to advantage in the design of systems of this type. Furthermore a considerable amount of work has been done in recent years by computer manufacturers as well as several academic groups toward developing interpretive digital computer programs which expedite the use of existing component analysis programs as subroutines in overall system programs. The objective of this work is to make it possible for engineers with only rudimentary knowledge of computer programming to use the machines. In a number of cases, programs have been developed which make it possible for an engineer to read into the computer with a system configuration and trial values for parameters expressed in common engineering language, and obtain from the computer, system outputs and other variables of interest which correspond to his assumptions.

Tables of analog computer programs for specific elements are also available in the literature of several fields. Although it is more difficult to systematize the use of an analog computer than a digital computer, an analog computer program is inherently patterned after the physical system. Thus it is relatively easy to investigate various element configurations and different parameter values once the element programs are available.

The use of computers in implementing system design procedures based on successive analysis of trial designs is well established and refined techniques are available in many fields. Present research is directed toward enlarging and extending the scope of element and subsystem analysis programs, and in decreasing the time and effort required by engineers in using interpretive programs.

b. Completely Automated Design Procedures: A new approach which uses a digital computer to automate the entire design procedure has been shown to be practical for special systems particularly in the area of control systems. This type of approach is receiving a considerable amount of research attention and is likely to grow in importance in the next decade.

The procedure referred to is based on machine techniques for handling optimization problems such as those which have been treated for many years through use of the calculus of variations. One form for a problem of this type is the following: Given (1) a set of ordinary differential equations describing certain unalterable elements of the system, and (2) a mathematical statement of the criterion of performance, find (1) that input to the unalterable elements which produces an extremum of the performance criterion subject to the restraint imposed by the unalterable elements, and (2) find a configuration of elements which produces this optimum input from the available variables (including the desired system inputs). Mathematically such problems can be expressed as follows. The unalterable elements can be described by a vector state equation of the form

$$\frac{d\bar{x}}{dt} = f(\bar{x}, \bar{y}, t), \quad \bar{x}(0) = \bar{c} \tag{14}$$

where $\bar{x}$ is the state vector, for the unalterable elements, $\bar{y}$ is a vector of inputs to the unalterable elements, and f is a real valued function of $\bar{x}$, $\bar{y}$ and t. The performance criterion can frequently be expressed as a functional of $\bar{y}$, $J(\bar{y})$, by equations such as

$$J(\bar{y}) = \int_0^T q(\bar{x}, \bar{y})\, dt \tag{15}$$

where q is an instantaneous error function which is integrated over an interval (0, T) to give the performance criterion. The design problem can now be stated as: (1) Determine $\bar{y}$ to minimize $J(\bar{y})$ subject to the restraint imposed by Equation (14), and (2) Determine a configuration of elements which will produce $\bar{y}$ from the desired system inputs and $\bar{x}$.

A less general, but sometimes more practical, formulation of the design problem eliminates the task of automatically determining a configuration of elements to produce $\bar{y}$ (item (2) above) by assuming that the configuration is given. In such a case the state equation becomes

$$\frac{d\bar{x}}{dt} = f(\bar{a}, \bar{x}, \bar{z}, t), \ \bar{x}(0) = \bar{c} \tag{16}$$

where $\bar{a}$ is a vector of system parameters which are to be determined by the design, and $\bar{z}$ is a vector of known system inputs. The performance criterion, J, now depends on the parameter $\bar{a}$, a finite set, rather than on the function $\bar{y}$ as in the previous formulation. The problem then becomes one of choosing the values of $\bar{a}$ to minimize J subject to the restriction of Equation (16).

The computational difficulties associated with either of these formulations of the design problem can be formidable. In particular the calculus of variations used alone does not usually yield the required numerical results. In the past ten years, however, several computational approaches have been developed which show general promise, and which have been shown to solve specific problems. These methods are: (1) Pontryagin's Maximum Principle (2) Bellman's Dynamic Programming, and (3) Techniques for implementing the Calculus of Variations type approach. Specific references to this work are given in the last section of the paper.

Example 2: This example is concerned with the design of an optimum control system using a procedure developed by Webb (see reference 41). An unalterable element satisfying the equation

$$\alpha\ddot{x} + \dot{x} = y \tag{17}$$

is given, where x is the output, and y is the input to the unalterable element, and α is a constant. A performance criterion J(y) given by

$$J(t) = \int_t^T [(x-x_i)^2 + \lambda^2 y^2]dt \tag{18}$$

is also specified, where x_i is the input to the system which is to be designed.

Use of Webb's design method results in the equations

$$\begin{aligned} \phi_{11} &= 1 - \phi_{12}^2 / \alpha^2\lambda^2 \\ \phi_{22} &= \phi_{22}^2 / \alpha^2\lambda^2 - 2\phi_{12} - 2/\alpha\ \phi_{22} \\ \phi_{12} &= -(\phi_{12}\phi_{22}) / \alpha^2\lambda^2 + \phi_{11} - \phi_{12}/\alpha \end{aligned} \tag{19}$$

which were solved on a digital computer for the special case of $\alpha = 1$, $\lambda = 1/4$. The structure of the final system is shown in Figure 8, and its responses to a step input for various choices of T are shown in Figure 9. Some further details of the solution are given in Appendix 2.

8. Use of Computers in Simulation of Systems

The present category cannot be entirely separated from previously discussed methods of using computers in system work. In particular the approach to be discussed in this section is closely related to the approach discussed in section 6 since the objective in either case is to determine system outputs for certain inputs and initial conditions. The distinction between the use of a computer as aid in implementing system analysis and its use as a simulator can be best appreciated in the light of several general remarks concerning models.

In the typical application, a model is chosen to simulate certain properties of a physical system in such a way that there is a one-to-one correspondence between the variables in the model and certain of the system variables. The way in which the variables in the model change with respect to changes in inputs, system structure, etc., can then be established from "experiments" with the model, and these results can be used to infer corresponding properties of the physical variables. Although mathematical models have emerged in recent years as extremely effective in dealing with systems from many branches of engineering, such models are by no means the only type. The scale model, used by architects and engineers in the mechanical sciences, is a well known example of an entirely different type of model.

A simulation of a system with a computer is a use of the computer to model the system. The computer elements and subroutines can usually be described in mathematical terms, and this mathematical description of the computer program for a system simulation is also a mathematical model for the system. The distinction between a computer model and the use of a computer in implementing calculations for a mathematical model is the following. In the use of a computer as a simulator available computer elements, subroutines, etc., are taken as given, and elements and operations from this restricted class are used to model the system. On the other hand, when the computer is used to implement calculations, the mathematical model is usually specified. Thus in one case, the computer approximates properties of the system directly; and in the other, the computer approximates the solution of equations which in turn approximate properties of the system.

Several examples of computer simulation in diverse fields will now be described, and a simulation problem from ceramic engineering will be formulated and solved in an Appendix.

Business Simulator: In the area of economics and business management, digital computer programs have been developed to combine the solution of supply and demand equations with a set of logical decisions to produce a model for the operation of a restricted segment of the business community. The computer inputs and outputs have been tailored so as to make it possible to use the simulator in management "games." In this problem the use of the computer to simulate the system rather than implement the solution of equations is clear cut. Mathematical models of the complexity needed to represent the system were not developed prior to stimulation experiments, since their use was impractical with hand calculations.

Monte Carlo Studies: An important capability of the digital computer is its ability to select numbers from specified random populations. For example, a digital computer could be programmed so that the number at a particular location in the machine changes randomly with time in such a way as to have a prescribed probability distribution function. Systems from a number of fields can be simulated by a set of equations having inputs or parameters which change randomly with time according to some specified distribution function. Such simulations, which are frequently called Monte Carlo studies, have been useful in modeling highway traffic flow, work sampling activity, and communication systems to mention several applications.

Dynamic System Simulation: Much effective work has been done in simulating dynamic systems on computers. The nature of such problems lends itself well to analog computers, and dynamic system simulation is very likely the most important area in which the analog computer is applied. The success of the analog computer in simulating dynamic systems lies in the fact that most such systems can be adequately modeled in terms of ideal elements, and these ideal elements in turn can be easily modeled by computer elements.

Some of the simpler ideal physical elements, the equations which they satisfy, and the analog computer program for their simulation are given in Table I. As an example of a direct simulation, consider the elementary system of Figure 10(a). A model for this system using computer simulation of the ideal elements is given in Figure 10(b).

In this example the computer circuits simulating the spring, the mass, and the friction were devised as the best analog computer approximation to the equations for the ideal mechanical elements. The simulation point of view, however, is to think of the analog computer program as a model for the physical system. Thus attention is focused on the accuracy with which voltage in the computer approximates velocity in the physical system rather than considering first the approximation to the physical system afforded by the mathematical model, and then the approximation to the mathematical model afforded by the analog computer elements. A great deal of work has been done in devising computer circuits to model physical elements. Tables and discussions of many of these circuits are given in the references of section 10.

Example 3: This example is concerned with the simulation of a mixing process and its associated control equipment on an analog computer. Consider the process shown in Figure 11. Liquid (or gases) from two sources - one at a high temperature T_b and another at a low temperature T_c - are combined and the composite material flows into a tank containing liquid at a uniform temperature, T. Liquid at temperature T flows out of the mixing tank. The temperature T is measured, and a controller is employed to maintain T at constant reference value, T_r, by automatically adjusting valves on the incoming hot and cold lines. The problem chosen for consideration is that of determining the error between T and T_r after a sudden disturbance in the temperature T_c. This error is a function of time and will depend on the settings of the proportional, derivative, and integral gains of the controller.

Use of the equations given in Appendix 3 results in the computer circuit of Figure 12. This circuit was used to plot the curves of Figures 13 and 14 which show, on a normalized basis, error vs time for various controller settings. The setting corresponding to $\bar{k}_p = 0.5$, $\bar{k}_d = 0.5$, $\bar{k}_i = 0$ gives the smallest transient error.

9. Computers as Elements of Process Control Systems

In section 1 of this paper it was pointed out that computers have two distinct roles in systems work, namely (1) computers can be used in system design, and (2) computers can be used as elements in process control systems. The use of computers in system design was covered in sections 4-8. The purpose of this section is to define and discuss briefly the role of computers as system elements.

In a certain sense the use of special purpose computers as control system elements began almost concurrently with the conception of control systems. An error detector which computes output minus input, a temperature sensor, and almost any other control system element can be thought of as a special purpose computer in the sense that the element output is a prescribed mathematical function of its inputs. The objective of this section will be limited to a consideration of the use of general purpose computers as elements in physical systems, but of course it is difficult to make a sharp distinction between special purpose and general purpose computers.

General purpose computers are incorporated into systems in three different ways.

(1) The computer can store information which is used to compute, on the basis of prescribed rules, the required time varying inputs to parts of a system. This is essentially an open loop mode of operation since the computer processes only stored data without sensing the actual state of the system. This use of the computer will thus be termed "open loop operation."

(2) The computer can also be operated from a fixed program to process both stored data and data from sensors coupled to the physical system. In such an application feedback around the computer is present, and this type of operation will be identified as "feedback operation."

(3) The most sophisticated use of a computer results when the computer program as well as certain data are determined by the state of the physical system. In such an application there is feedback around the computer, and the potential to change the computer program makes the system adaptive. Thus this type of operation will be termed "feedback-adaptive." Several characteristics of systems using computers in these three ways will now be discussed.

a. General: Several problems and characteristics which are common to all, or at least two, of the three types of operation are the following.

Use of the digital computer with physical elements poses an immediate problem. Most physical variables change continuously with time while on the other hand most digital computers are variables which change in discrete steps. Thus almost all applications of a digital computer as an element of a physical system require analog-to-digital or digital-to-analog converters or both.

Applications (2) and (3), which have feedback around the computer, can be unstable so that design for stable operation becomes one of the more serious problems. The same two applications raise problems of correctly identifying the current state of the physical system, and choosing an appropriate measure of system performance.

As a final general comment, any use of a general purpose computer as a system element requires economic justification. It is clear that in industrial applications the savings in actual cost of the product of a system must be large in order to justify the use of expensive computing equipment.

b. Open Loop Operation: A system using a computer in an open loop mode is frequently termed a computer directed system. A block diagram of this type of operation is given in Figure 15, where it should be noted that the flow of information is entirely from the computer memory through the computational operations to the input of the physical components.

One well publicized application of the use of a computer in the open loop mode is in the numerical control of machine tools. For such an application the dimensions of a desired product are stored in the computer memory and used to compute the desired position of a cutting tool at any time. This information is then used to control the machine.

The use of a computer for open loop operation is the most elementary of the cases being considered. The absence of feedback removes instability as one of the design problems, but accuracy of the controlled variable cannot be as high as for the cases where feedback is used. The advantage provided by the computer is flexibility in being able to perform complicated operations on the stored input data to produce the necessary input for the physical system.

c. Feedback Operation: Use of a computer in this mode makes it possible to sense the current state of a physical process and use this information along with stored information to determine the required input to the physical elements. A block diagram for this type of operation is shown in Figure 16. The added feedback around the computer makes stability a serious problem, but provides a more accurate means of controlling the desired variables. Note that in this case the input to the physical elements is computed from the stored data and from measurements of the current state of the system. The added information provided in this mode of operation makes it possible to correct errors as they occur, rather than operating with a fixed procedure. However, the computer program is fixed in advance so that the functional relation between computer inputs and outputs does not change.

A flight control system for a modern space vehicle affords an illustration of this use of a computer. In such a case an equation specifying the optimum trajectory from any point in space to the desired target is programmed into an on-board computer. The computer then uses a measurement of its current position along with stored target data as inputs from which to compute the required thrust vector at all points of the powered flight.

d. Feedback-adaptive operation: This type of operation has the greatest potential of the three types being discussed. As indicated in Figure 17, the computer accepts inputs from stored data, from sensors coupled to the physical process, and also from other sources. In this case the program for the computer can be altered by its inputs so that the system can adapt to changes in inventory, economic factors such as sales price of the product, or any other input which is pertinent and which can be quantitatively related to the desired output of the system.

Among serious design problems presented by this very general type of operation are choice of criteria for weighting various factors which influence the desired system outputs, and accurate means for identifying the current state of the physical process.

Examples of the use of computers in systems occurring in the ceramics industry will be given in other papers which are a part of this symposium.

10. References

This section lists references which cover in detail the areas summarized in the paper. The references are divided into the following general categories to conform with existing patterns in the literature:

1. General - references which treat both analog and digital computers.
2. Characteristics and Uses of Analog Computers.
3. Simulation using Analog Computers.
4. Analog computers as System Elements.
5. Numerical Methods - this is a mathematical area useful in preparing problems for digital computers.
6. Characteristics of Digital Computers.
7. Use of Digital Computers in System Design by Successive Analysis.
8. Use of Digital Computers in Optimum Design Problems.
9. Use of Digital Computers in Simulation.
10. Digital Computers as System Elements.
11. References Pertaining to Specific Examples.

General

1. E. M. Grabbe, Editor, Automation in Business and Industry, John Wiley and Sons, 1957.

2. P. von Handel, Electronic Computers: Fundamentals, Systems, Applications, Prentice Hall, 1961.

Characteristics and Uses of Analog Computers

3. A. E. Rogers and T. W. Connolly, Analog Computation in Engineering Design, McGraw Hill, 1960.

4. C. L. Johnson, Analog Computer Techniques, 2nd Edition, McGraw Hill, 1963.

5. A. S. Jackson, Analog Computation, McGraw Hill, 1960.

6. S. Fifer, Analog Computation, Vols. 1-4, McGraw Hill, 1961.

Simulation Using Analog Computers

7. V. L. Larrowe, "Direct Simulation-Bypasses Mathematics-Simplifies Analysis," Control Engineering, Nov. 1954.

8. J. M. Carteron, "Analog Representation of Heat Exchange. Application to the Simulation of Heat Exchangers of Nuclear Power Plants," Proceedings of the Computers in Control Systems Conference, AIEE, May 1958.

9. W. J. Karplus, Analog Simulation, McGraw Hill, 1958. This book includes an extensive bibliography to 1957.

10. C. W. Worley, "Simulation of Industrial Processes with the Analog Computer," Automatic Control, p. 44, Jan. 1959.

11. D. E. N. Davies, "Electronic Simulation and Computer Techniques in the Design of Automatic Control Systems," Journal of the British Institution of Radio Engineers, p. 323, April, 1961.

12. F. A. Woods, "Simulation of Process Control with an Analog Computer," Ind. Eng. Chem. 50, p. 1627-30, November 1958.

Analog Computers as Systems Elements

13. L. F. Kazda, "Control System Optimization Using Computers as Control System Elements," Proceedings of the Computers in Control Systems Conference, AIEE, May, 1958.

14. R. J. Bibbers, "Process Regulation with Analogue Controls," Automatic Control, Nov. 1954.

15. E. T. Davis, "An Electronic Control System for Industrial Control Applications," Instruments, June 1951.

16. J. G. Truxal, "Computers in Automatic Control Systems," IRE Proc., p. 305, Jan. 1961.

Numerical Methods

17. R. W. Hamming, Numerical Methods for Scientists and Engineers, McGraw Hill, 1962.

18. F. B. Hilderbrand, Introduction to Numerical Analysis, McGraw Hill, 1956.

19. R. Bellman, Dynamic Programming, Princeton University Press, 1957.

Characteristics of Digital Computers

20. R. Nathan, E. Hanes, Computer Programming Handbook, A Guide for Beginners, Prentice-Hall, 1962.

21. I. Flores, Computer Logic, Prentice-Hall, 1960.

22. Computer Issue, Proceedings IRE, Vol. 41, No. 10, October 1953.

23. Computer Issue, Proceedings IRE, Vol. 49, No. 1, January 1961.

Use of Digital Computers in System Design by Successive Analysis

24. B. A. Bryce, "Circuit Synthesizing Techniques with the Aid of a Digital Computer: An Annotated Review of the Literature to 1962," Document 278 635, July 16, 1962. This report is a review of 161 references pertinent to direct use of digital computers in design, and other topics.

Use of Digital Computers in Optimum Design Problems

25. R. Bellman, Adaptive Control Processes: A Guided Tour, Princeton University Press, 1961. This book contains an extensive bibliography.

26. L. S. Pontryagin, et al., The Mathematical Theory of Optimal Processes, John Wiley, 1962.

27. R. Bellman, et al., "A Computational Procedure for Optimal System Design and Utilization."

Use of Digital Computers in Simulation

28. IBM, "General Purpose Systems Simulator," Reference Manual for a System simulation program published by IBM Corp.

29. J. Moshman, "The Generation of Pseudo-random Numbers on a Decimal Machine," J. Assoc. Computing Mach., Vol. 1, p. 88-91, 1954.

30. M. Palevsky, "An Approach to Digital Simulation," Proc. National Simulation Conference, Jan. 1956.

31. F. I. Musk, "A Monte Carlo Simulation of a Production Planning Problem," Computer J. Vol. 2, p. 80-94, July 1959.

32. C. P. Bonini, "A Simulation of a Business Firm," Proceeding of the Joint Computer Conference, Vol. 21, 1962.

Digital Computer as System Elements

33. M. Pelegrin, et al., "Digital Computers in Automatic Control," Proceedings of the Computers in Control Systems Conference, AIEE, May 1958.

34. E. L. Braun, "Digital Computers in Continuous Control Systems," IRE Trans. Vol. EC-7 p. 123-128, June 1958.

35. A. K. Susskind, "Digital Information Process for Machine Tool Control," IRE Trans. Vol. EC-7, p. 136-140, June 1958.

References Pertaining to Specific Examples

Example 1

36. M. F. Berard, Diffusion in Ceramic Sys: A Selected Bibliography, Iowa State Univ. Aimes Lab. Res. & Development Rpt. USAEC IS-448.

37. G. E. Forsythe, P. C. Rosenbloom, Numerical Analysis & Partial Diff. Eq., John Wiley, 1958, survey of field.

38. L. W. Ehrlick, "A Numerical Method of Solving a Heat Flow Problem with Moving Boundary," Asso. for Comp. Mach. Vol. 5, 1958, p. 161-176.

39. J. Crank, The Mathematics of Diffusion, Oxford University Press, London, 1956.

40. G. H. Bruce, et al., "Calculation s of Unsteady-State Gas Flow through a Porous Media," Petroleum Trans. A.I.M.E. 198(1953) pp. 79-92.

Example 2

41. R. P. Webb, "Synthesis of Measurement Systems," Ph.D. Thesis, Georgia Institute of Technology, July, 1963.

Appendix* 1: Further Discussion of Example 1

As discussed in the body of this paper, Frick's law given as Equation 12 is the basic relation which determines the concentration of diffused substance as a function of time and distance. The normalized variables defined on page 12 reduce Equation 12 to the equation

$$\frac{\partial C}{\partial T} = \frac{\partial}{\partial X}\left[D^*(C)\,\frac{\partial C}{\partial X}\right] \qquad A_1 - 1$$

For the particular D(C) being used in this example, a variable s, defined by

$$s = \int_0^C DdC \Big/ \int_0^1 DdC,$$

is introduced so that Equation A-1 becomes

$$\frac{\partial s}{\partial T} = D^*(C)\,\frac{\partial^2 s}{\partial X^2} \qquad A_1 - 2$$

with the boundary conditions

$$s = 1,\ X = \pm 1 \quad \text{for } T > 0.$$
$$s = 0,\ -1 < X < 1 \text{ for } T = 0.$$

To solve Equation A-2 on a digital computer, it must be reduced to a difference equation. The technique used in this example is known as the Crank-Nicolson method (see reference 39, p. 189) and in the case being considered it results in a set of equations of the form

$$s(m,n+1) = s(m,n) + \frac{\delta T}{2(\delta X)^2}\Big[9s(m,n+1) + 9s(m,n) + 2\Big]\Big[s(m+1,n+1) + s(m+1,n) - 2s(m,n+1) - 2s(m,n) + s(m-1,n+1) + s(m-1,n)\Big] \qquad A_1 - 3$$

where m is the distance index, n is the time index, δT is the step size in time and δX is the step size in distance. The special equation

$$s(o,n+1) = s(o,n) + \frac{\delta T}{2(\delta X)^2}\Big[9s(o,n+1) + 9s(o,n) + 2\Big]\Big[s(1,n+1) + s(1,n) - s(o,n+1) - s(o,n)\Big] \qquad A_1 - 4$$

is required for $m = 0$. Equations $A_1 - 3$ and $A_1 - 4$ were solved for 8 steps in X (i.e. m = 0, 1,2,3,4,5,6,7) making $\delta X = 0.125$. A ratio $\delta T/(\delta X)^2 = 1/10$, which makes $\delta T = 0.0016$, was chosen to insure convergence of the method used.

*The assistance of A. W. Bowers in programming this example is greatfully acknowledged.

Equations $A_1 - 3$ and $A_1 - 4$ became a set of 8 simultaneous nonlinear algebraic equations for any fixed n. The procedure used in going from step n in T to n + 1 was the following. The simultaneous equations were made linear by assuming a value for one of the s(m,n) in every term involving a product of two s terms. The set of linear equations was then solved for the s(m,n) and the solutions compared to the assumed values. The new values were then taken as a better approximation and the procedure was iterated until the assumed solutions and actual solutions agreed within an error in the seventh decimal place. This completed a solution at the stage n + 1 and the answers for the s(m,n) as well as for concentration were printed out.

The next step in time (i.e. to n + 2) was taken by assuming as an initial guess

$$s(m,n+1) = s(m,n+2)$$

and proceeding as before. Very small steps in T were required to insure convergence of the iterative process. Further the initial concentration stated in Equation $A_1 - 2$ was not satisfactory as a starting point. The values given by Crank (see reference 39) at T = 0.1724 were used to start the process and approximately 300 iterations were taken to produce the curves of Figure 7. (Crank worked the example through 5 iterations by hand.)

The computing time required on the Borroughs B500 was several minutes.

Appendix 2: Further Discussion of Example 2

Webb's procedure (see reference 41) makes possible the design of a control system as follows. An equation, in this case Equation 17, which describes unalterable elements of the system must be given. In addition a performance criterion such as expressed by Equation 18 must be specified. The performance criterion must be in the form of the integral of a certain instantaneous error - in this case the squared difference between the instantaneous control system input and output plus a constant times the squared input to the unalterable elements. The latter term measures the energy supplied to the unalterable elements.

By using a modified dynamic programming approach, Webb obtains an expression for an input to the unalterable elements which minimizes the error functional expressed in this case by Equation 18. The expression for this input, y_{min}, gives y_{min} as a function of the system input, the output state variables, defined in this case as $x-x_i$ and $d/dt(x-x_i)$, and the time variable parameters $\emptyset_{11}(t)$, $\emptyset_{12}(t)$ and $\emptyset_{22}(t)$. In this case the latter parameters are the solution of Equation 19. Equation 19 is in general a set of nonlinear differential equations which in most cases must be solved on a digital computer. Once the Ø's are known, the expression for y_{min} determines a system configuration, such as that of Figure 8, which will generate y_{min} from the output state variables and the system input. The exact nature of the system depends on the length of the time interval, T, over which the system is designed to minimize the performance functional. If T is chosen to be long (i.e. $T = \infty$), then the Ø's are constant and a linear time invariant system results. If T is not long then a time variable system results.

Appendix 3: Further Discussion of Example 3

The components of the mixing process shown in Figure 11 are described mathematically as follows. The temperature T in the mixing tank is assumed to satisfy the equation

$$dT/dt + \frac{m}{M} T = \frac{1}{M}(m_h T_h + m_c T_c) \qquad A_3 - 1$$

where m is the mass flowing out of the tank at temperature T, m_L and m_c are the mass flows at T_h and T_c, and M is the mass of liquid in the tank. The valves are assumed to be described by an equation such that m_h and m_c are related to the controller output, V, by the equations

$$m_h = k \int V\, dt$$

$$A_3 - 2$$

$$m_c = m - k \int V\, dt\ .$$

The controller satisfies the equation

$$V(t) = k_p(T-T_r) + k_d \frac{d}{dt}(T-T_r) + k_i \int (T-T_r)dt. \qquad A_3 - 3$$

Combining Equations $A_3 - 1$ and $A_3 - 2$, yields

$$dT/dt + \frac{m}{M} T = \frac{m}{M} T_c + \frac{k}{M}(T_h - T_c) \int V(t)\, dt \qquad A_3 - 4$$

The following choice of constants and definitions were made

$$m/M = 0.1$$
$$\bar{k}_p = \frac{k}{M}(T_h - T_c)k_p$$
$$\bar{k}_d = \frac{k}{M}(T_h - T_c)k_d$$
$$\bar{k}_i = \frac{k}{M}(T_h - T_c)k_i$$

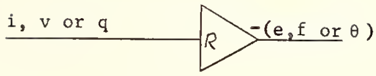
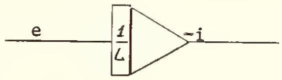
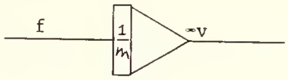
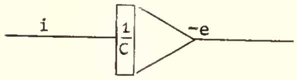
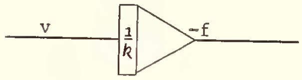
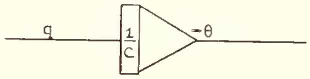
Since the system is linear, no loss in generality is incurred by taking $T_r = 0$, and this results in the equation

$$\text{Error} = T(t) - T_r = T(t) \qquad A_3 - 5$$

The response in error due to a sudden change in T_c was investigated, and this was produced mathematically by letting T_c be an impulse at t = 0 which is equivalent in this case to an initial condition on T.

The computer diagram which solves Equations $A_3 - 3$ and $A_3 - 4$ subject to the assumptions listed is given in Figure 12. Plots of T(t)/T(t = 0) for various choices for $\bar{k}_p$, $\bar{k}_d$, and $\bar{k}_i$ are given in Figures 13 and 14.

Table 1

Description of Element	Equation	Analog Simulation
Electrical resistance	$e = Ri$	
Viscous friction	$f = Rv$	
Thermal resistance	$\theta = Rq$	
Electrical Inductance	$e = L\ di/dt$	
Mechanical mass	$f = m\ dv/dt$	
Electrical capacitance	$e = \frac{1}{c} \int idt$	
Mechanical spring	$f = \frac{1}{k} \int vdt$	
Thermal mass	$\theta = \frac{1}{c} \int qdt$	

Glossary of Symbols:

e - voltage; f - force; θ - temperature
i - current; v - velocity; q - heat flow
R - resistance in all systems
L - electrical inductance
m - mechanical mass
c - electrical or thermal capacitance
1/k- mechanical spring constant

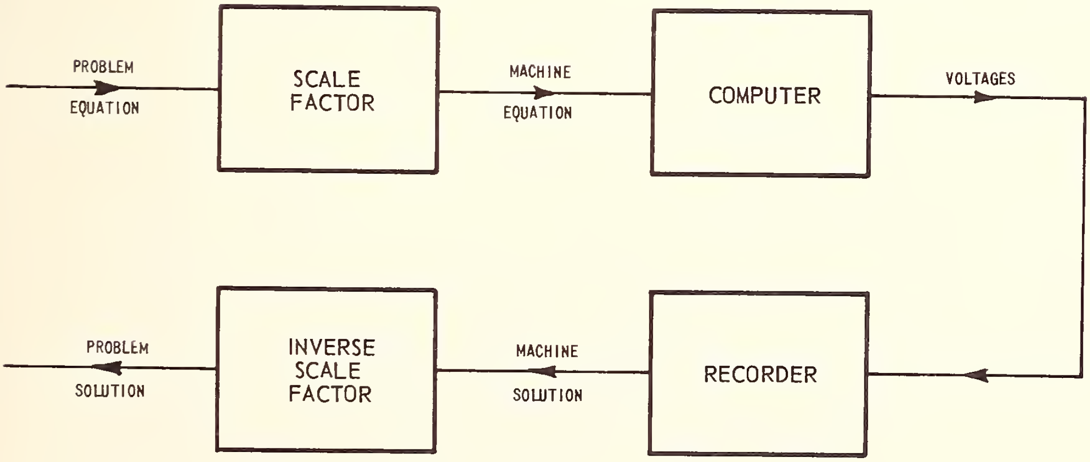


Figure 1. Basic Steps in Using an Analog Computer.

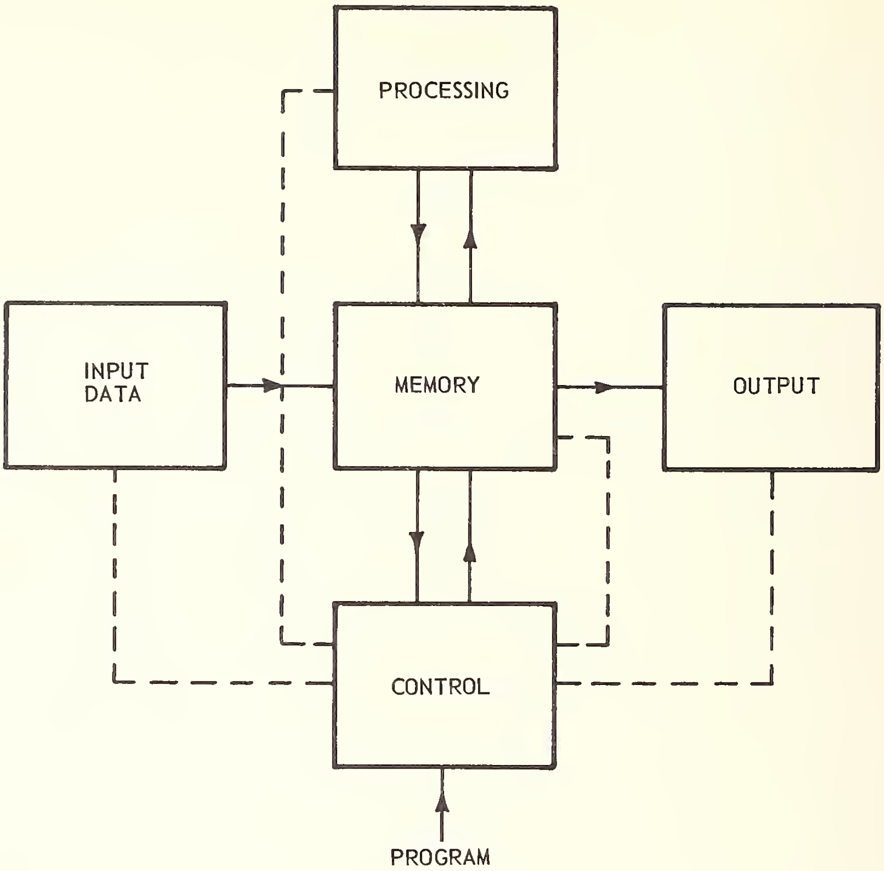


Figure 2. Basic Steps Performed by a Digital Computer.

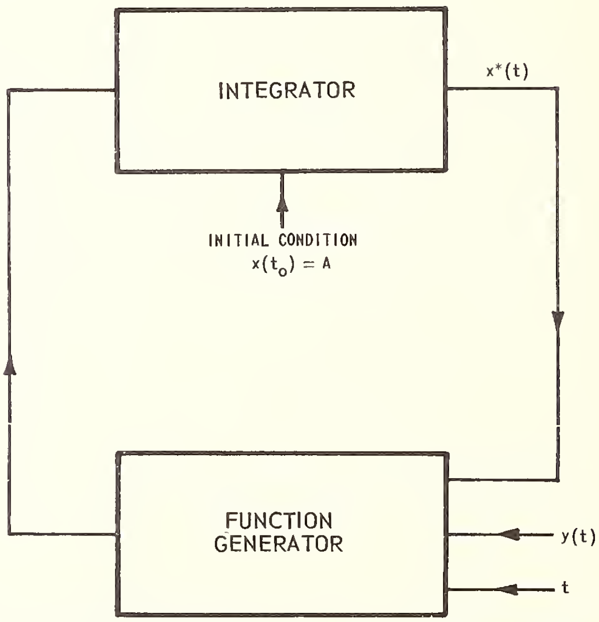


Figure 3. Approximating the Solution of a Differential Equation with an Analog Computer.

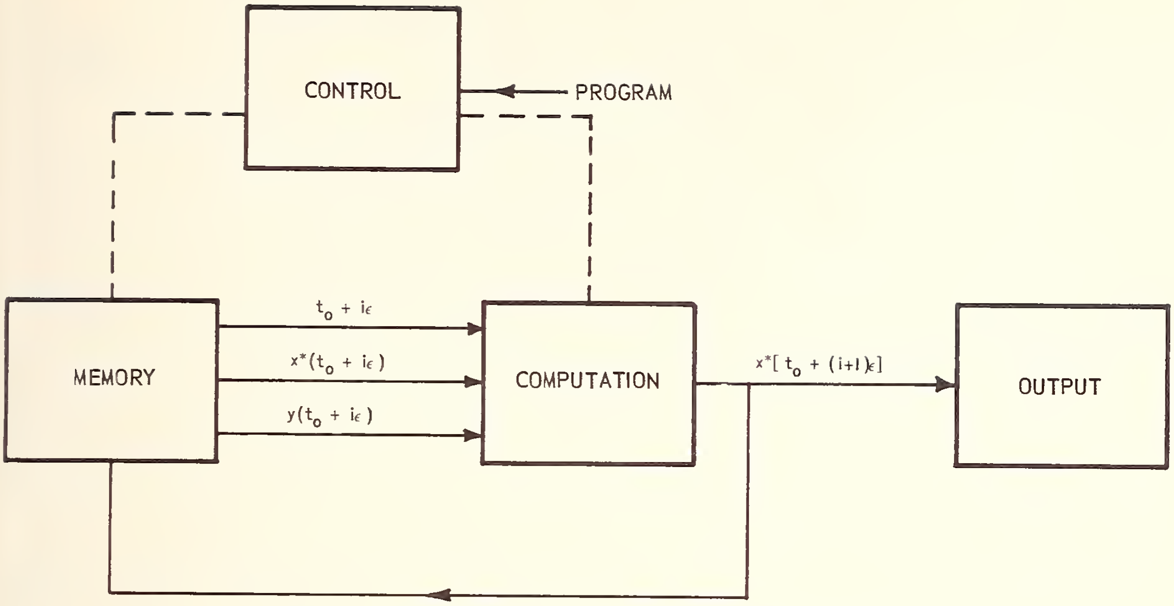


Figure 4. Approximating the Solution of a Differential Equation with a Digital Computer.

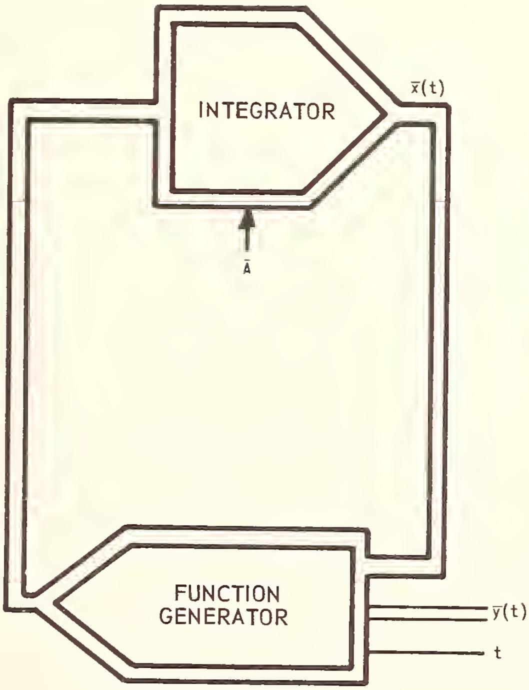


Figure 5. Logical Program for Solving a Set of Simultaneous Differential Equations.

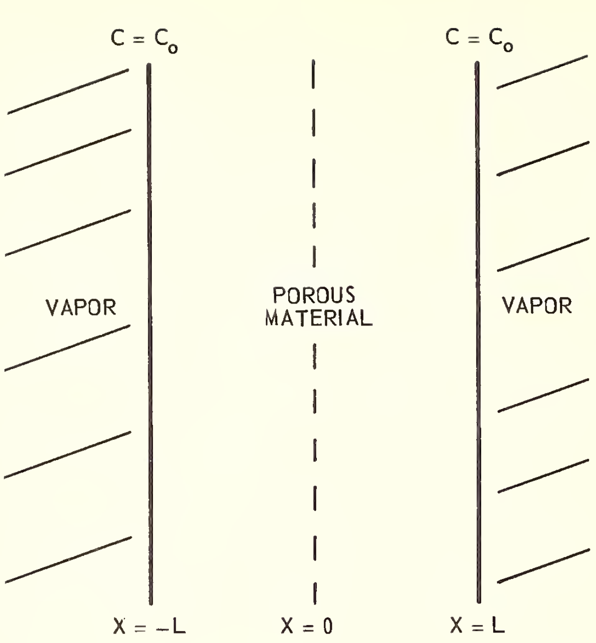


Figure 6. An Illustrative Diffusion Problem.

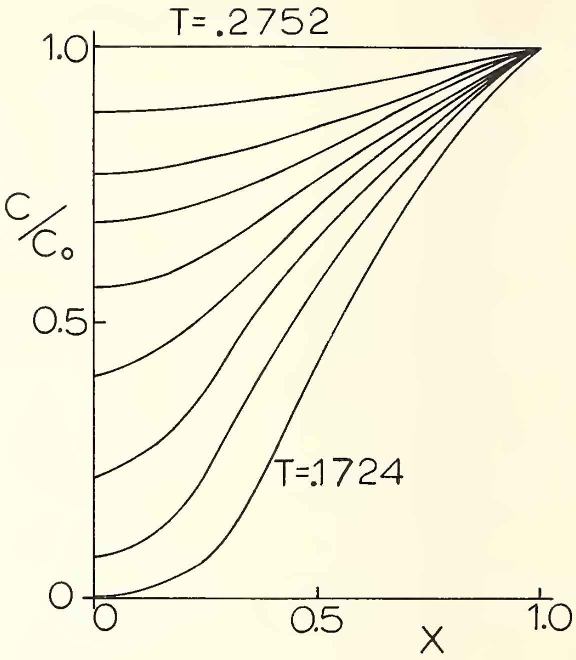


Figure 7. Curves Illustrating the Solution of a Diffusion Problem.

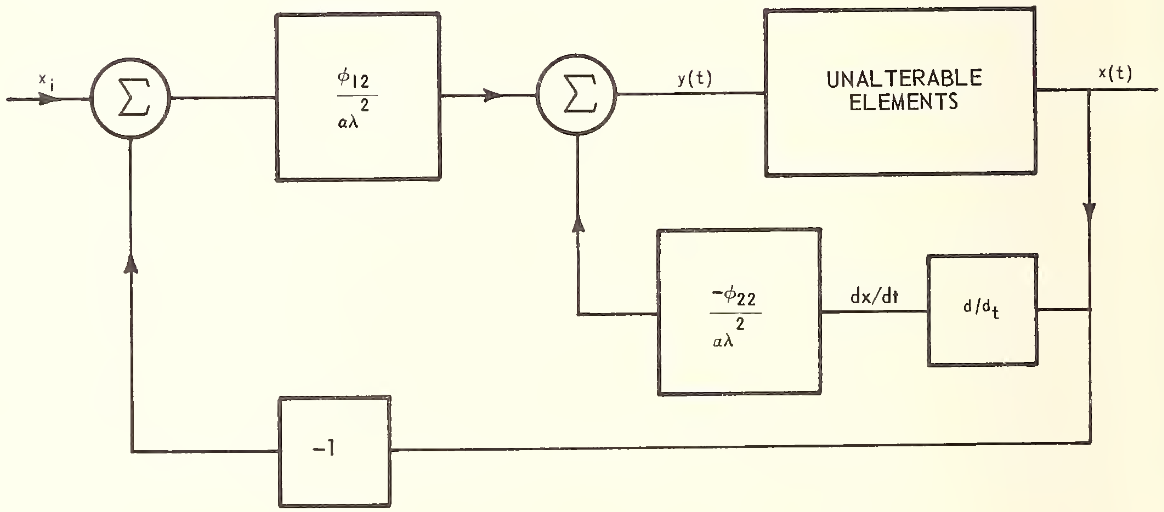


Figure 8. System Configuration for a Control System.

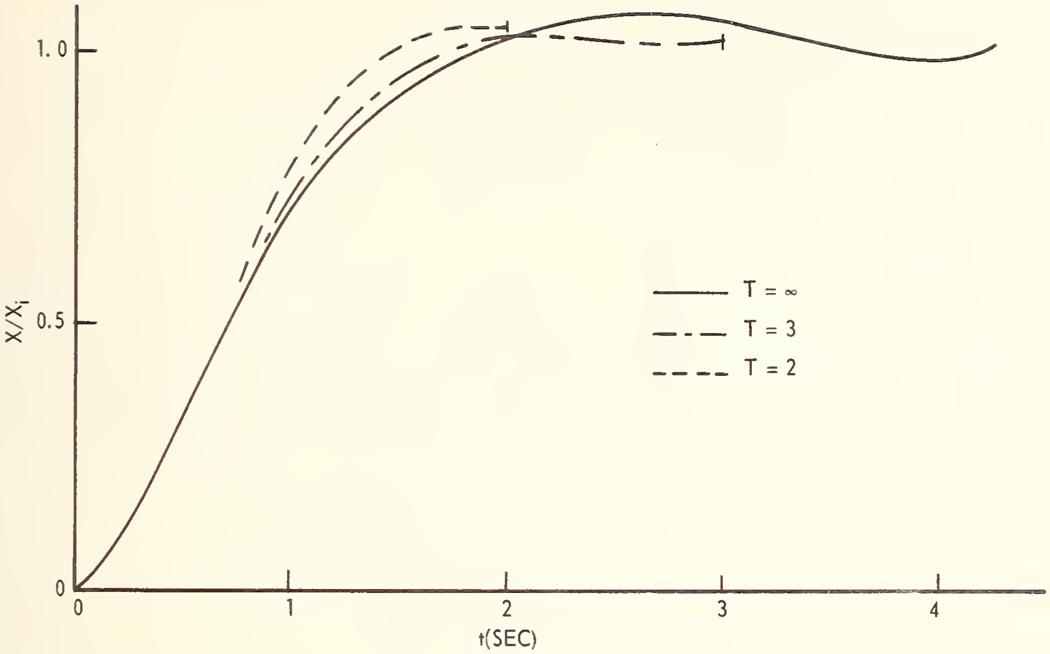


Figure 9. Step Responses of a Control System.

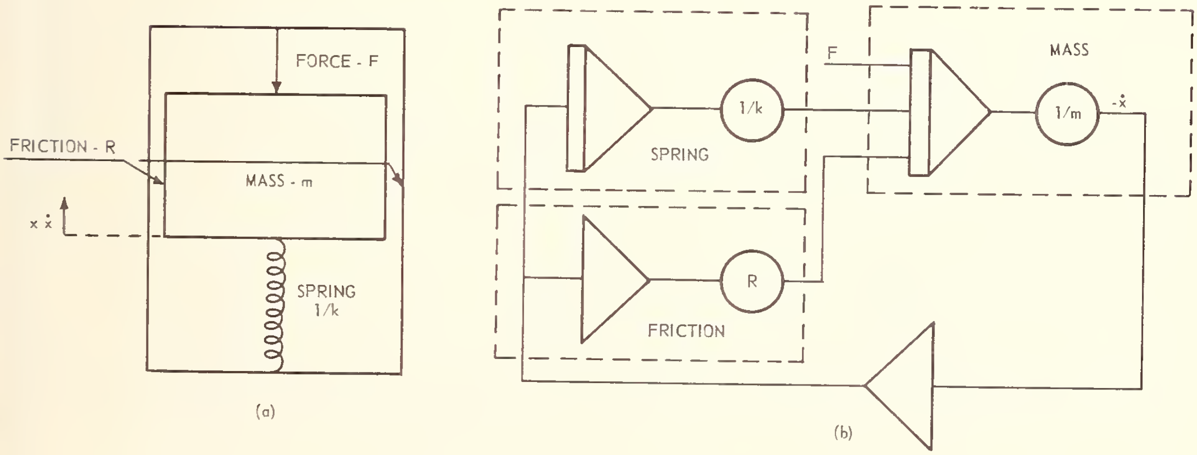


Figure 10. A Simple System and its Simulation on an Analog Computer.

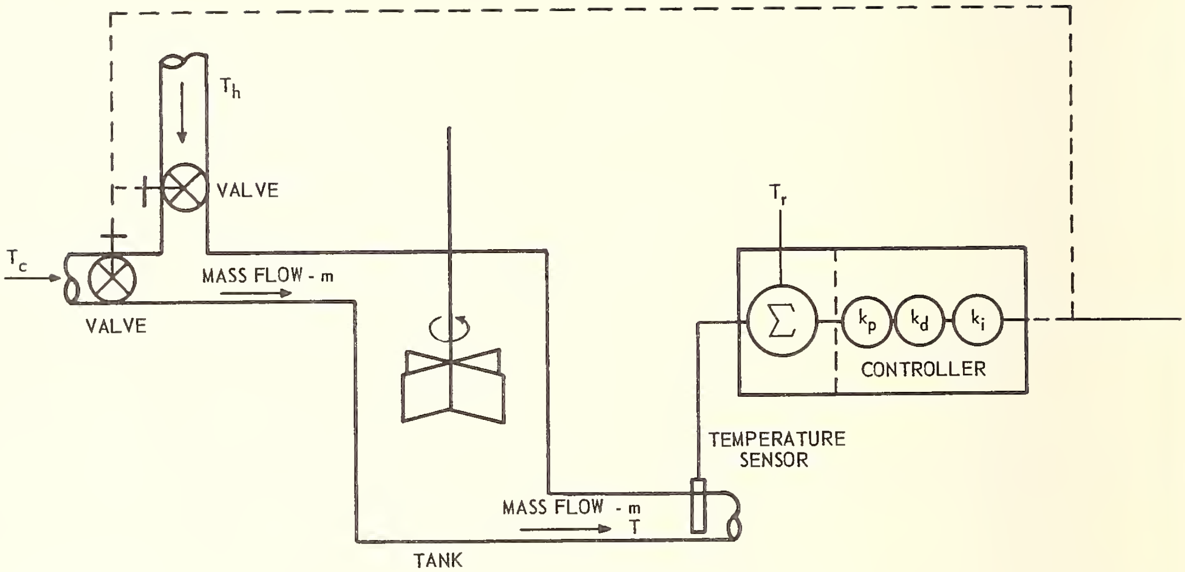


Figure 11. A Mixing Process.

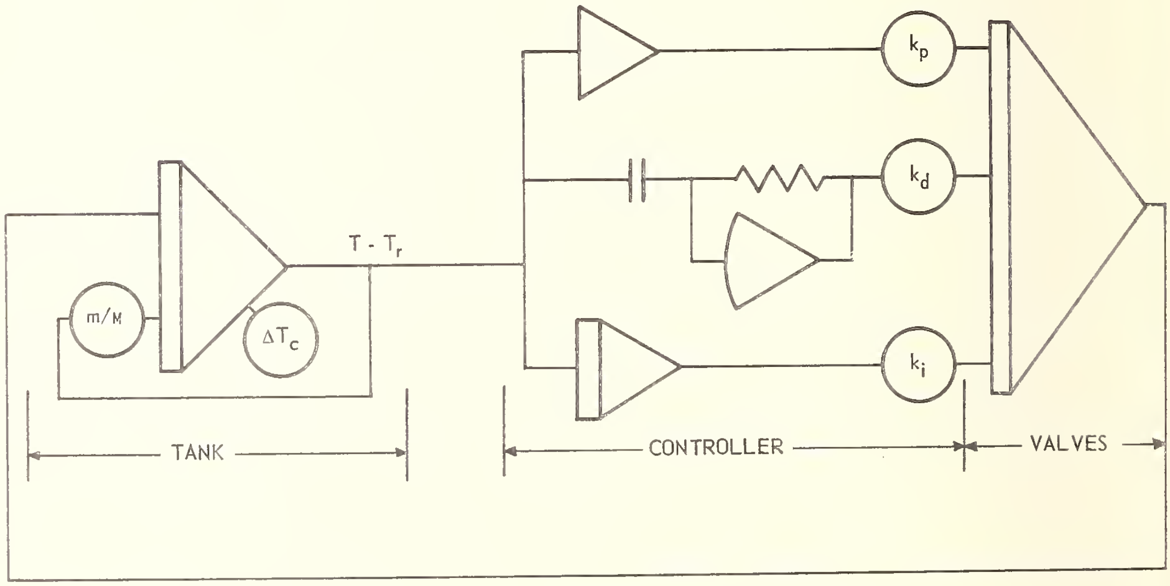


Figure 12. Computer Simulation for a Mixing Process.

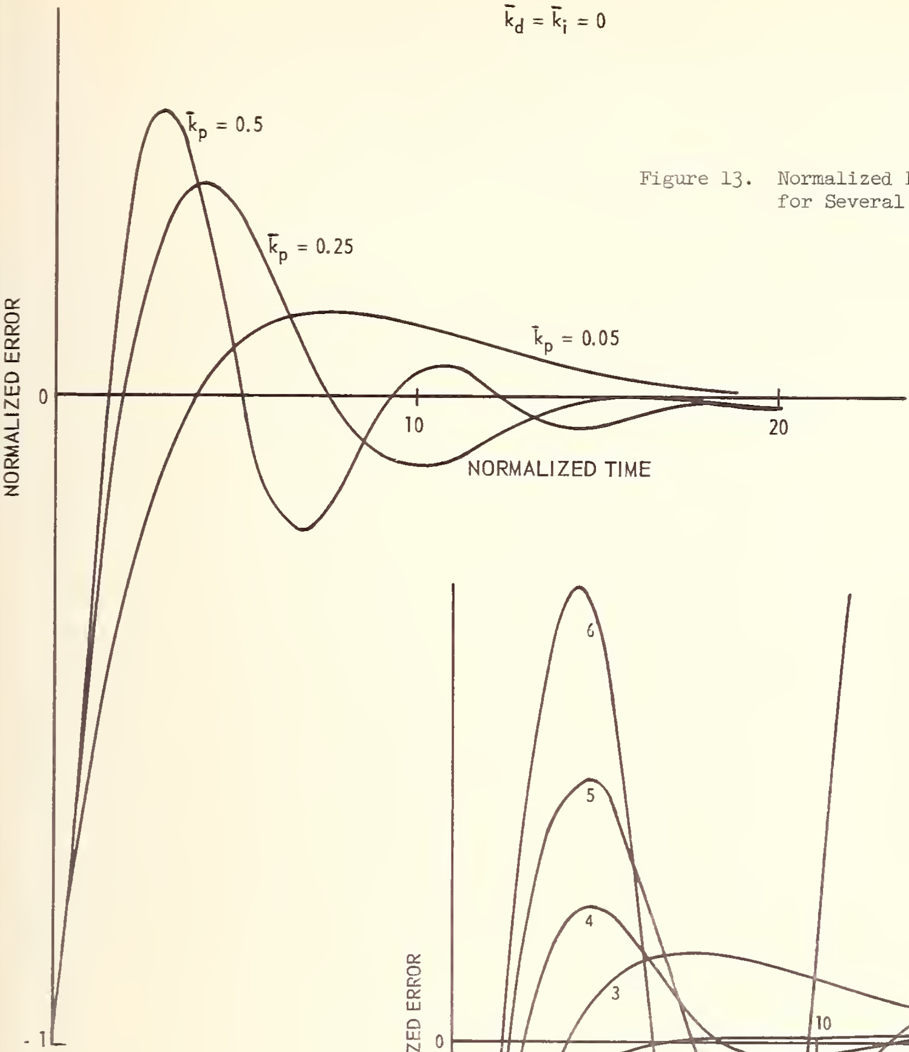


Figure 13. Normalized Error vs Normalized Time for Several Gain Settings.

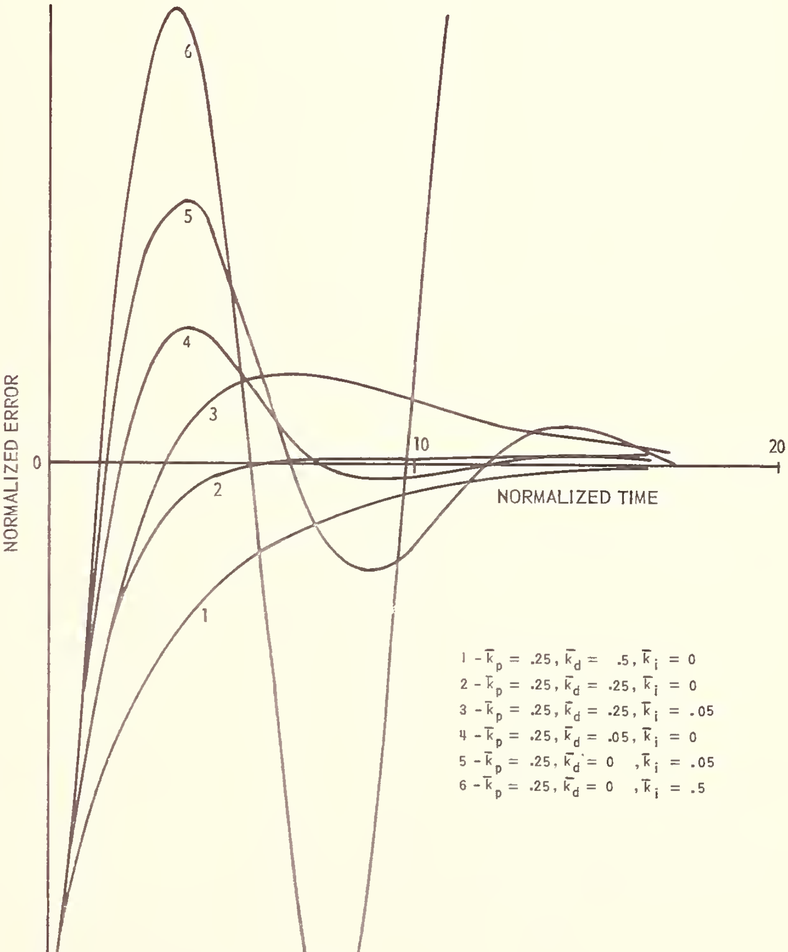


Figure 14. Normalized Error vs Normalized Time for Several Settings of $\bar{k}_p$, $\bar{k}_d$ and $\bar{k}_i$.

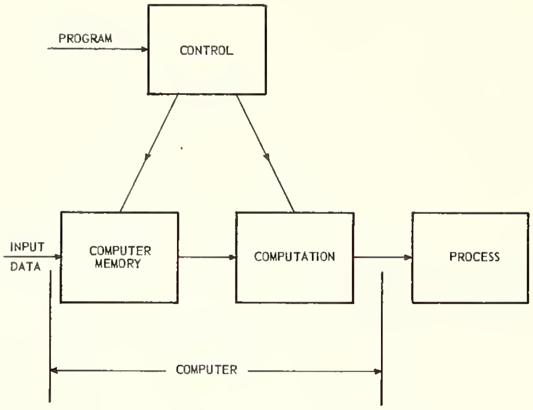
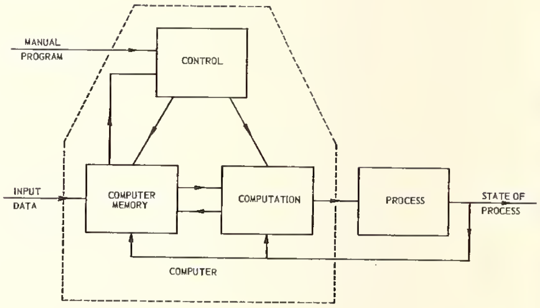


Figure 15. Block Diagram Showing Open Loop Computer Operation.

Figure 16. Block Diagram Showing Feedback Operation.



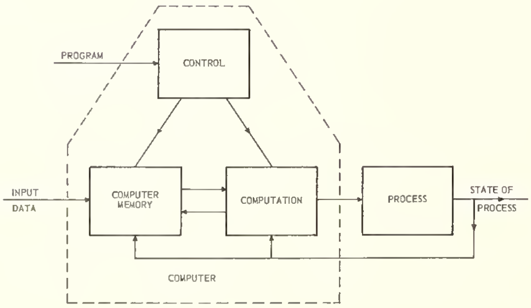


Figure 17. Block Diagram Showing Feedback-Adaptive Operation.

Statistical Methods in Systems Design

W. C. Hackler

1. Introduction

When a manufacturing process can be characterized by means of a mathematical equation one is in a position of greater understanding, hence greater ability to exercise control over the process. Too often in the ceramic industries qualitative relationships are understood but the appropriate constants associated with a mathematical function of the process variables are frequently unknown. Statistical methodology provides a means of estimating the necessary coefficients to associate with each process variable, thus providing the needed quantitative relationship between variables and process yield or quality. In addition it provides a level of assurance (probability) depending on the magnitude of the fluctuations or variability encountered in the process. Such fluctuations may be assignable to raw materials, equipment, or level of operational variables, temperature for example. In any system, decisions need to be made at various stages throughout the process as to whether the process is performing in a satisfactory manner or not. In order that a valid decision be made, one must have assurance that test data is a reliabile predictor of performance. Statistical methods are most useful in reaching a decision - either adjust the process or allow it to continue. In many instances it has been proved that frequent adjustment of some process variable was the most upsetting influence. Some variation about an average value is to be expected and as long as this variability is within specified limits, the system should not be adjusted for these minor fluctuations. Another contribution of statistical methods is that of evaluation of "noise" or error in the system. Techniques are available whereby the masking effect of one variable can be isolated so as to get a true measure of the desired variable. Error analysis and reduction of error (minimization of variability) are effectively handled by statistics as well as the determination of what is a minimum error for a particular system.

2. Estimation

A manufacturing process may be examined as a situation in which some response (yield, purity, strength, resistance or some other property of a product) is a direct result of the operational level of many variables (raw materials, equipment and process variables.) If the variables can change along a continuous scale and the response can be measured in the same manner, then it is possible to describe the response as some functional relationship between the variables. In many instances it has been found that an expanded Taylor series[1] will provide a suitable estimate of this response. The order of the series may be extended as required for a satisfactory fit. Such a relationship for two variables of the second order might be:

$$Y = b_{\cdot} + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2 + b_{11} x_1^2 + b_{22} x_2^2 \qquad (1)$$

When:

Y = response

x_i = i^{th} variable

b_i = differential effect associated with i^{th} variable

The order of the series depends on the complexity of the true but unknown surface. The magnitude of each $\underline{b}$ indicates the contribution of each variable to the response of interest. Hence, the variables with the largest $\underline{b}$ values are the ones that need closest control so as to maintain the response at the desired level.

The coefficient b_{12} in equation (1) is of prime concern. This term x_1x_2 is known as the interaction term or the failure of x_1 to produce the same effect on the response at different levels of x_2.

Interaction can best be explained by examination of Figure 1. When a response is plotted against a variable x_1 for two different levels of a second variable x_2, two lines result, one for the low value of x_2 and one for the higher level of x_2. If these two lines are parallel then the effect of the variables on response is additive, and it may be assumed that there is no interaction between the two variables. When there is lack of parallelism, interaction between x_1 and x_2 exists. Then the effect of x_2 on the response is dependent on the level of x_1. Hence interaction may be characterized as factor dependence insofar as the response is concerned. Inspection of the set of numbers in Table I is also a helpful way of evaluating interaction.

Table I

x_1 \ x_2	Low	High
Low	5	10
High	9	18

Increase of x_1 from the low to the high value increased the yield by 4 units. An increase in x_2 from the low to the high value increase the yield by 5 units. Simultaneous increase of x_1 and x_2 from low to high value resulted in an increase in yield of 13 units. Since 13 does not equal 5 + 4, interaction is present. Thus interaction between two variables can be plotted or set up in a tabular form. When the number of variables increases, the ability to "see" the interaction diminishes. However, they can be estimated in a quantitative manner by means of the functional relationship (equation 1) regardless of the number of variables.

A most useful outcome of equation (1) is sometimes referred to as response surface methodology since the equation describes a "surface" of some order. The plot in figure 2 is just like a topographic map where points of equal elevation over a given area are connected by a contour line. These contours locate all combinations of x_1 and x_2 within a prescribed region that will yield a given response. This contour system points out the necessity of simultaneous evaluation of the effect of both variables on the response. A given response is not achieved in a unique manner rather there are many combinations of variables which will yield the desired level.

Another interesting feature of the contour plot is the necessary control which must be maintained over each variable. It is quite apparent that the solution to this problem depends on the slope of the contours in the operational region. In figure 2 a deviation of _a_ units of x_2 would produce much more drastic results when operating at point _R_ as compared to point _S_. The closeness of spacing of the contours has the same implication in this plot as does a topographic map. Close spacing of contours is indicative of a steep slope or a rapid change in response. In this case it would be much better to choose point _S_ as operating levels of the variables to give a response of 80 rather than point _R_, since the slope is steeper at _R_ than it is at _S_.

It is quite probable that more than one property of a product must be kept within desired limits. The response surface for a second property will be a quite different relationship than that determined for the first. Such an example is seen in figure 3. Let the contours labeled 60, 70, and 80 represent yields and the second family of contours labeled _A_, _B_, and _C_ represent cost per unit, increasing from _A_ to _C_. Since maximum yield is desired, anywhere within the 80 contour would meet this requirement. However, in order to maximize yield and minimize costs, a smaller operational region results. The segment of the _A_ curve between points _M_ and _N_ represents conditions where a yield of at least 80 is obtained with minimum cost.

Graphical interpretation can be extended to three dimensions by stacking planes such as seen in figure 2 to yield figure 4. In figure 4 it can be seen that the contours are the intercepts of a surface with a plane at specified levels of the third variable.

The functional relationship of equation(1) may be compared to a transfer function relating a response to the differential contribution of certain operational variables. This equation is one which is empirically determined, by methods of least squares so as to result in the best fit of the data. Data can come from laboratory experiments, pilot plant trials, or best of all from production runs. Often times "evolutionary operations" techniques serve as a useful tool for collecting data from production runs so as to minimize upsetting the process. Such information will characterize the response surface, pointing out optimum operational combinations of controlled variables. The effect of control of each variable is readily available from this equation. Such a relationship provides a method of comparing one design or control system against an alternative one. Thus a quantitative method is available for evaluation of a particular design.

In the estimation of the function of equation (1) there will be some variability in response which cannot be associated with an assignable cause or variable. For example, the failure of duplicate observations to be identical is such a source of error. The contribution of an uncontrolled variable results in a large increase of the error term. Any variability which cannot be accounted for is pooled in the error term - sometimes called lack of fit. As the number of uncontrolled variables increases, more specifically as the error term increases, the usefulness of the equation diminishes. One has to describe what is a tolerable error term, based on product specifications and the economics involved. In some cases the error cannot be reduced to a desired level without major changes in the equipment, control systems, or the processing itself. In any event a measure of the "noise" in the system is available as lack of fit of the function relationship. The magnitude of this error serves as "yardstick" for evaluating the effect of specified variables. If the contribution of a given variable is of the same order of magnitude as is the error, then one cannot say that the specific variable contributes to the response. If, on the other hand, the effect of a variable is much greater than the magnitude of the error, it can be safely assumed that the particular variable is a critical one. Thus as the magnitude of the "noise" or error term diminishes the ability to discriminate between the effects of variables increases. This error term provides the necessary scale for the evaluation of alternative designs.

3. Raw Materials

Before any system can be set up in a satisfactory manner some control over the starting materials must be established. Whether material comes from a variety of suppliers or whether material is mined locally, there is an associated inherent variability. In order to be assured that the system has the opportunity to perform at its optimum, then realistic material specifications (based on past performance or from pilot plant studies) need to be maintained. Too frequently specifications are established for no good reason, leading to inadequate or too stringent requirement for the material. It is essential that quality measurements made on the starting materials be a reasonable predictor of the performance of this material in the process. Thus, if measurements of sensitive properties of the material are not obtained control may appear to be had when in actuality there is no correlation between measurements and performance. Figure 5 portrays such a situation of low correlation between measurements and outcome in the process. Once such a plot is made it is quite apparent that the measurement will not predict the outcome. One could just as well "fit" line A as line B to this set of data since really there is not a satisfactory linear fit to the data.

One method of evaluation of the quality of raw materials is by utilization of an $\bar{X}$ and R chart for measurements[2] on one or more sensitive properties of the material. By use of the $\bar{X}$ chart the level of a variable can be kept under surveillance. The R chart, however, is a measure of variability within a sample, thus it is an estimate of the variability in the lot of material from which the samples were taken. Figures 6 and 7 depict the $\bar{X}$ and R charts respectively.

The $\bar{X}$ chart will present the level of quality of a given sample of material while the R chart is indicative of the variability within this sample. Two points are to be remembered here:

1. These charts as shown are prepared for averages of samples of a given size. If the sample size changes, the upper and lower control limits will also change. An increased sample size implies a better estimate of the lot average of the particular property being determined, hence control limits move nearer to the average. Consequently, the control limits for the $\bar{X}$ chart in figure 6 are not appropriate when considering individual items.

2. A second aspect of these charts is that they present estimates of the property of the population from which they were taken. As such these estimates are totally independent of desired values or specifications. In the event that the specifications fall outside the control limits for individual items, then one readily accepts the material being evaluated. However, when the specifications fall inside the control limits for individuals, the material does not meet the specifications. It is quite possible that materials or parts could be prepared under controlled conditions and indicate a state of statistical control, yet not be able to meet specifications. When this situation arises there are at least three alternatives available: 1) find new vendors, 2) help vendor improve his process, 3) examine specifications to see if they are realistic.

There are many cases in which incoming material is evaluated as being either acceptable or not acceptable. The absolute value of a characteristic is not measured. All that is known is that the inspected item meets or fails to meet specifications. It is not known how good or how bad an item may be. Also a part could be rejected for any one of a number of reasons such as size, smoothness, glaze defects or warpage. When a sample from a lot of material or parts is inspected in such a manner, this is known as sampling by attributes. Items such as resistors, capacitors, wall tile, dinnerware, and enameling stock are examples of items which could be inspected in such a manner.

To help in the classification of incoming material the United States government has developed sampling methods published as MIL-STD-105-D.[3] These plans take into consideration the lot size and acceptable quality level. Based on this information sampling plans are specified for single, double, and multiple samples. The sample size is specified and the number of defective items allowable in a good sample are recorded. An excellent feature of MIL-STD-105-D is the fact that the operating characteristic curve for each plan is presented.

The operating characteristic (OC) curve is a measure of the ability of a given sampling plan to discriminate between acceptable and unacceptable material. Such a curve is shown in figure 8.

If the lot size is at least ten times greater than the size of the sample, these curves can be obtained from the summation of the appropriate terms of the binomial expansion $(q + p)^n$ when q = fraction acceptable
p = fraction defective
n = sample size.

In the case of the curve labeled N = 150, c = 8 the sample size is 150 and the c refers to the number allowable defective items in the sample. At a given fraction defective (0.04 for example) the probability of acceptance is obtained from:

$$P_r = C_r^n (0.96)^{n-r} \; (.0.4)^r$$

where r indicates the number of defective items. The probability is obtained by summing probabilities for no defectives, one defective, etc. up to eight defectives at that specified value of fraction defective. When $p < 0.10$ and $pn < 5$, a Poisson distribution[2] is easier to use to compute P. Such data can be read directly from a chart for the Poisson distribution found in Dodge and Romig.[4] It can be seen that the curve N = 150, C = 8 has a steeper slope, hence is more critical. The probability of accepting a sample of p = 0.04 is about 0.80 while the other curve will accept such a lot approximately 0.06. Lots of p = 0.08 will be accepted only 20% of the time by N = 150, c = 8 as compared to 43% of the time for N = 150, c = 11. Thus the sample size and value of c may be found for any desired probability of acceptance at a specified fraction defective. As a general rule a probability of 0.95 is satisfactory for the acceptance of a good lot. This remaining 5% is known as the producers risk (α) the risk of producer takes of having a good lot rejected. By the same token the risk of accepting a poor lot is known as the consumers' risk (β).

A sampling plan is characterized by any two of the three following:

a) fraction defective
b) producer's risk
c) consumer's risk

There are a variety of alternatives that may require some compromise so as to come near to satisfying all three conditions. The manner in which a OC curve changes with sample size is illustrated in figure 9. As the sample size increases the plan becomes more discriminatory in evaluating good and bad material. Naturally the ideal curve is a shaped one. As sample size increases the curve approaches that shape.

4. Process Control

Statistical techniques are employed throughout a process in order to detect shifts in the process as quickly as possible so that adjustments can be made to return the process to its desired operating conditions. The adjustments may come in the form of manual changes to a piece of equipment by an operator or it may be in the form of an electrical feedback from a sensing element to the piece of equipment. Some decision making element such as a computer is needed in the line as seen in figure 10. In this case two sensings are made,[5] furnace temperature and nominal core resistance. This information is sent to a computer which solves the transfer function with the data supplied by these sensings and feeds the correction to the furnace controls. It is to be remembered that again a correlation between the quality of the product and the process measurement must exist if effective control is to be maintained. Quite frequently the sensing is made of something other than the product or semi-finished product (e.g. temperature, atmosphere, or flow rate of the product). This avoids handling or contact with the product, but a good correlation must exist. Frequently too, it is the combination of two or more process variables (e.g. time and temperature) which must be sensed and controlled simultaneously. The importance of this simultaneous control is indicated by the magnitude of the coefficients of the interaction terms of the response function (equation 1).

An interesting example of the use of control of an automatic process was reported by Boatwright.[6] The process was that of making carbon coated resistors where control of the furnace temperature, rate of flow of gas, and speed of cone travel through the furnace was maintained. An X chart was plotted by a computer and use was made of a non-parametric test to evaluate shifts in the process average. Averages of five resistance measurements were plotted on a chart with the desired nominal resistance as the central value. The value of the average range was programmed as an estimate of the anticipated value based on past experience. Significant trends were established by any of the following conditions:

1. any point falling outside the control limits
2. two out of three successive points in zone A or beyond
3. four out of five successive points in zone B or beyond
4. eight successive points in zone C or beyond

A control chart is shown in figure 11. The width of a zone is equivalent to one standard deviation. Corrections for average values of resistance were calculated by the computer, the correction being based on an average of 20 or 30 previous points, not on the point at which a decision was made.

5. Reliability

To many people reliability of a given item is achieved by successive inspections and measurements, discarding unsatisfactory pieces and retaining the good. Such a process is merely sorting, making the best of what is available rather than making a better product. Reliability is used in many senses but the most appropriate connotation is that of a probability[7] of satisfactory performance over some specified period of time. As such it is neither a quality nor a property of an item, rather it is a number between 0 and 1 with no units. Thus reliability may refer to a system or a single component in a complex system. It is often evaluated in a negative sense, such as a failure rate or mean time between failures.

Since reliability is concerned with the ability of a given product to perform in a specific environment, it is apparent that an intermediate heat duty fireclay refractory would not have the same reliability as a 70% alumina refractory if the environment was near the service limit for the 70% alumina refractory. However, reliability of both refractories would be nearer to each other when service conditions were near those appropriate to an intermediate heat duty fireclay. Hence, it is assumed that material be selected wisely for a specified environment. The technique of "derating" is often used to increase reliability. If design indicates that a one watt resistor is an appropriate component then the use of a two watt resistor gives assurance or acts as a safety factor to increase reliability of the part selected.

It is quite possible that a given process can be producing a product under excellent conditions of statistical control yet the product might have a low reliability. This would indicate the presence of unknown or uncontrolled variables in the process which exerted a significant effect on the product. Apparently specification as well as processing control need be changed in order to characterize the product more properly as pertains to performance in service. Such a situation readily indicates the difference between the quality control function and the reliability function. Where the responsibility of quality control is present time, the responsibility of the reliability function is in the future.

Before an estimate of reliability of a product can be determined it is essential that tests be made. By and large, extended periods of time are required to obtain life test data, particularly when dealing with a product of low failure rate. Often life test data is obtained by means of accelerated tests, or by testing a large number of items for short periods of time. The equivalence of time and number of parts tested to yield the number of hours to first failure is not always justifiable. For a constant failure rate from the beginning of the test, this compression of time by increasing the number of components tested is on safer ground. Accelerated tests are often performed at higher stress levels than a part may be expected to see in service. In order to translate such test data to service conditions an appropriate factor must be known which will relate this overload endurance to expected life.

Reliability seen in a product is a direct result of reliability in the performance of a system which manufactures this product. The quality control function and the product reliability function must be brought into close proximity in order to achieve reliability in the system. System reliability is brought about by an understanding of the effect of process variables on the performance of the end product. Thus the design of the system - the materials used, the equipment selected, the closeness of controls, the monitoring of the performance - yields the level of reliability seen in the system. The major contribution of statistical methods is that of providing quantitative information relative to optimum operating levels of process variables. Increased complexity of systems seems to be the order of the day, and statistical design appears to be a most reasonable method to evaluate simultaneously the effect of many variables. In most systems the effect of one variable is not independent of another variable, hence the contribution of the interaction between two or more variables must be known. This is best accomplished by means of statistical design.

REFERENCES

1. W. G. Cochran and Gertrude M. Cox, Experimental Designs, Chapter 8A, John Wiley and Sons, Inc. (1957).

2. A. J. Duncan, Quality Control and Industrial Statistics, Richard D. Irwin, Inc. Homewood, Ill. (1959).

3. Sampling Procedures and Tables for Inspection by Attributes, MIL-STD-105D U. S. Government Printing Office, 29 April 1963.

4. H. F. Dodge and H. G. Romig, Sampling Inspection Tables - Single and Double Sampling, John Wiley and Sons, Inc. (1944).

5. M. D. Clark, H. K. Jordan, and R. C. Ward, "Precision Carbon Coatings by Computer Control," The Western Electric Engineer, 6 (2) April 21-27, (1962).

6. J. H. Boatright, "Computer-Plotted Control Charts," The Western Electric Engineer, 6 (1) January 20-21, (1962).

7. E.G.D. Paterson, "Reliability and Quality in Communications", Industrial Quality Control 20 (9) March 15-23, 31-38 (1964).

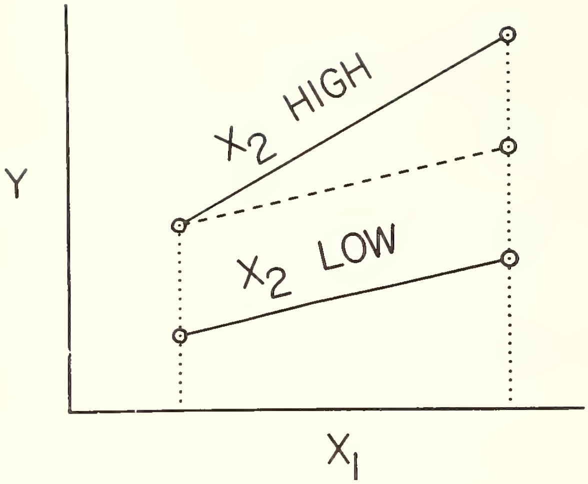


Figure 1. Interaction between two variables.

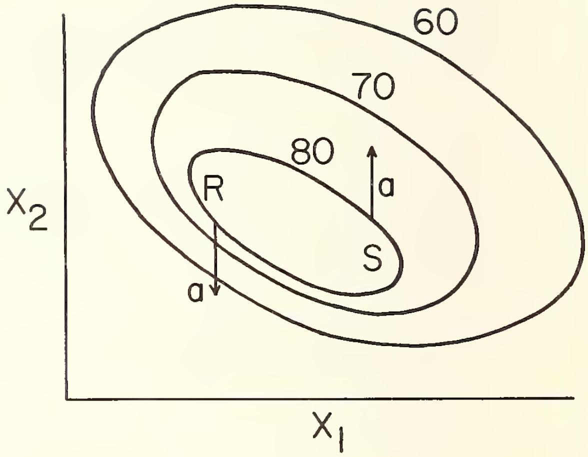


Figure 2. Contours of equal response in a two variable plane.

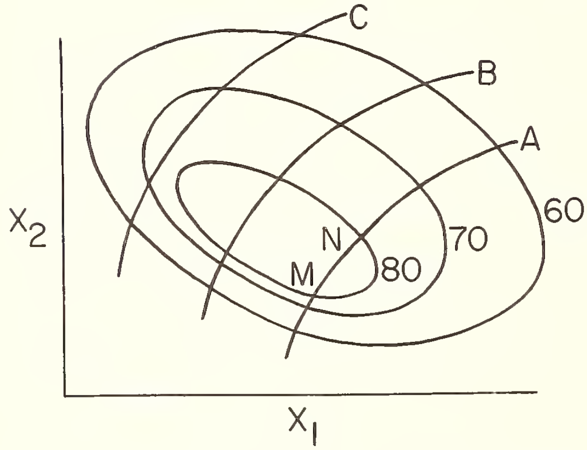


Figure 3. Families of contours for two different responses (60, 70, and 80) indicate % yield and A, B, and C indicate cost increasing from A to C.

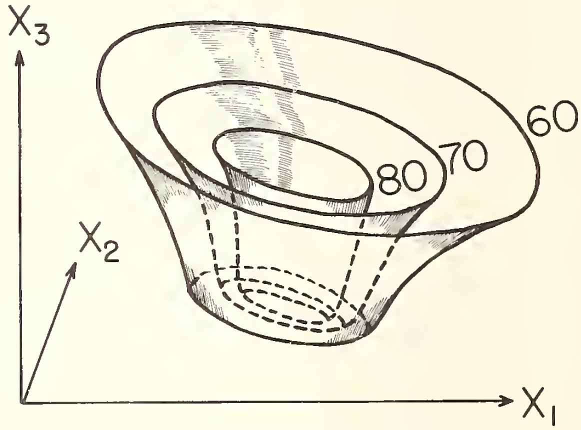


Figure 4. Response surface in X_1, X_2, and X_3 space.

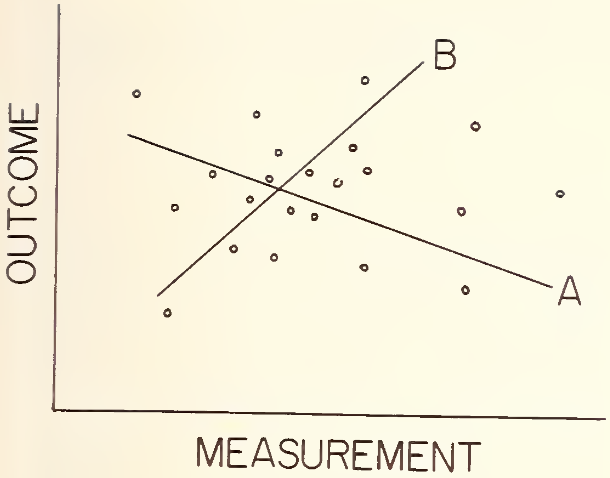


Figure 5. Lack of correlation between measured property and outcome or quality of a product.

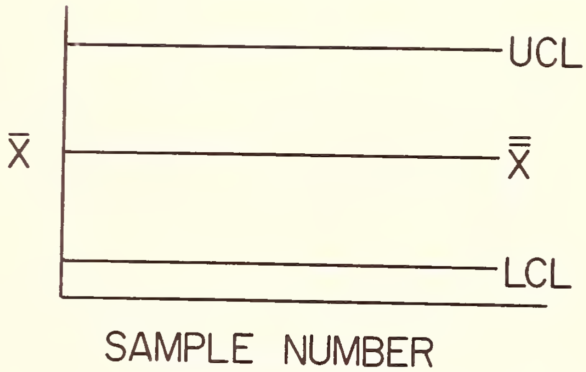


Figure 6. Control chart for sample average.

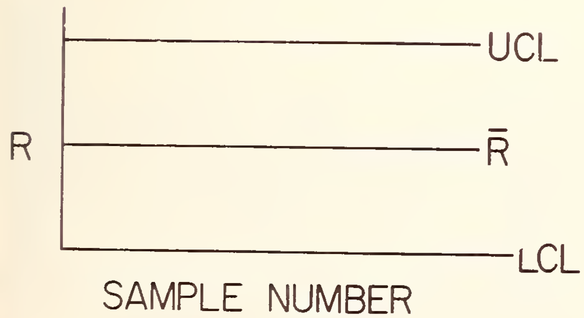


Figure 7. Control chart for sample range.

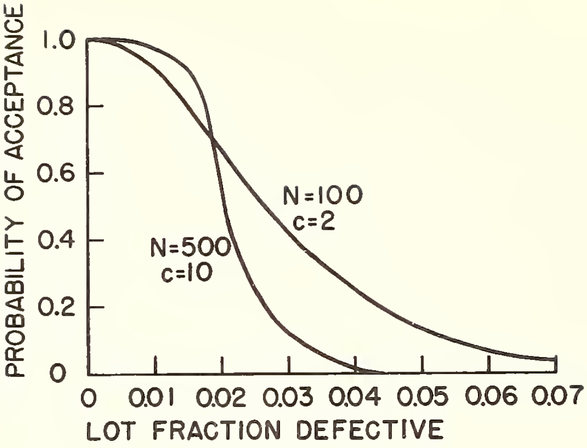


Figure 9. Operating characteristic curves for different sample sizes.

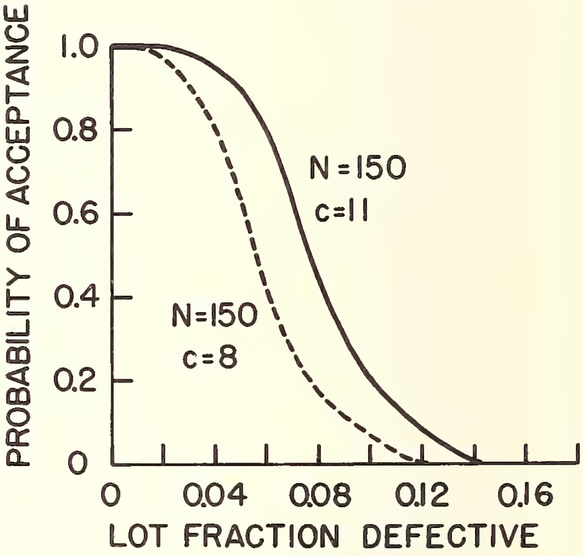


Figure 8. Operating characteristic curves for different acceptance numbers of defective items.

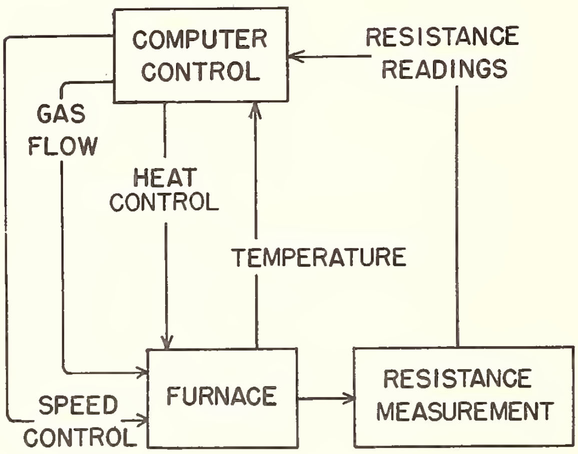


Figure 10. Computer controlled process.

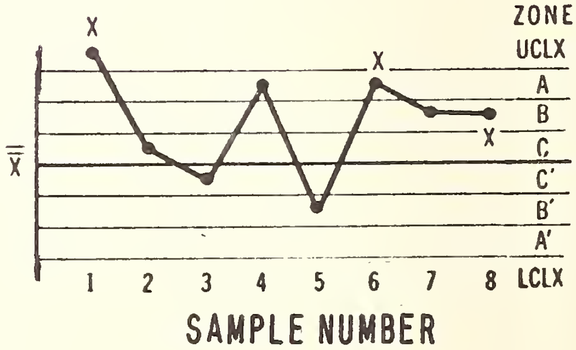


Figure 11. Control chart used in non-parametric test (after Boatwright).

Systems of Interest to Ceramic Engineers

A. H. Hurst & E. R. Bullock

1. Introduction

Beginning with the pottery of early civilizations, the study of ceramics has evolved into a highly scientific technology covering a broad range of endeavors. This evolution has been caused by the demands of modern industries which has placed an ever increasing premium on a more exact composition, precise dimension, or repeatable physical characteristic. In order to keep pace with these demands, the modern ceramist must rely on control technology to assist him in producing a product which will meet or exceed such stringent requirements.

In this paper, the authors describe some typical systems that are of interest to ceramic engineers. The examples were chosen in order to give a good cross section of the various types of systems available to industry. The practical aspects have been emphasized in order to give its reader a basic understanding of the mechanisms at work, and to stimulate his imagination toward applying the concepts to solve everyday problems. Both conventional analog control and more modern digital control systems are explored.

2. A Packaging System

Many small ceramic parts must be manufactured to extremely close tolerances and therefore have a significant cost which makes an accurate accounting of each part worth the investment of a system to count each part. This can be accomplished most easily during the packing phase of the operation when a particular number of parts are to be packaged in a container.

Figure 1 shows a simplified version of such a system which has as its basic parts a conveyor, a deflector bar, a chute, a photopickup, and an electronic preset counter. As parts move along the conveyor, they pass the photo pickup sensing head. The resultant electrical signal generated by the interruption of the light beam causes a single count to be registered in the electronic counter. At the same time, the deflector arm, capable of assuming two positions, is in position "A" in order that all parts will enter the container in position beneath chute "A". A reserve container is in position "B" so that when the container in position "A" is full, the delivery of parts can be directed immediately to "B", the full container moved out of the loader, and an empty container placed in position "A" ready to recieve parts when "B" becomes full.

The counter such as that illustrated in Figure 2 is capable of being preset to a given number which represents the desired quantity of parts to fill a container. When this number is reached, a relay closure is initiated within the counter and the counter immediately resets to zero in preparation of another cycle. A solenoid positions the deflection bar to "A" and "B" depending upon whether it is activated or not which in turn depends upon the alternating condition of the counter output relay.

While this example represents a simple open-loop control system, it can be readily adapted to other applications through variations in sensors or conveyors. Many parts are loaded by vibratory means which is less cumbersome and quite fast making it an attractive method for parts handling. Other sensors could be used depending on the nature of the parts to be loaded. Air sensors operate on the interruption of an air stream between a source and pressure diaphragm sensor, generating switch closures each time that the air pressure drops below a previously adjusted limit. This method would be useful with ceramic parts, but might lack the desired speed since the switch closure time could be in the order of milliseconds while the photo pickup is in the microsecond range.

A variable reluctance sensor would work only with magnetic materials such as ferroceramics, but has the advantage that it is low cost and quite dependable. This pickup consists of a permanent magnet surrounded by a coil. As the magnetic field surrounding the coil is disturbed by the presence of a magnetic material, the coil sees a magnetic flux change which generates an output signal in accordance with Lenz' Law. Other suitable variations will occur to the reader when faced with a specific packaging problem.

3. An Automatic Testing System

Quality assurance becomes increasingly more important daily. Materials and components must meet continually increasing demands on reliability which in turn requires greater knowledge of a larger number of components. To have a technician individually check each part is time consuming and therefore costly. A large amount of data can be obtained automatically with a system such as the one shown in Figure 3.

In this case, it is desired to know the temperature coefficient of ceramic dielectric materials. An oven with a large number of electrodes suitable for holding ceramic samples comprises the test jig. Connected to this is a commutator capable of selecting any of the samples and connecting it to a test oscillator. The test oscillator is running at a frequency dependent upon a known inductance-capacitance combination. The dielectric under test is connected in parallel with the tuned circuit capacitor, and thus causes a shift in frequency which is related to the value of capacitance of the sample.

A frequency counter measures the frequency of the test oscillator. This is accomplished by counting the cycles of the oscillator for an accurately known time, usually one second or decimal multiple thereof. At the end of the counting interval or "gate time", the counter input is closed, and the counting chain then holds a number which represents the number of cycles counted. If the gating interval were one second, the number would indicate cycles per second directly. The use of an electronic frequency counter allows the data to be measured, and converted to digital data at the same time. The resultant digital information can then be connected to the computer input directly.

A heater for the oven is controlled from the computer with on-off control through a relay. The temperature control loop is closed with a digital temperature transducer/encoder which converts temperature to a digital signal for direct input into the computer. The transducer/encoder such as that shown in Figure 4 consists of a mercury filled thermal bulb system which is connected to a bourdon tube. As the temperature increases, the pressure increases thus causing an expansion of the bourdon tube which results in rotary motion. The rotary motion, which usually moves a pointer on a scale, in this case rotates a shaft position encoder.

The shaft position encoder graphically shown in Figure 5 consists of a digitally coded glass disc with patterns of black and clear such that light can pass through some areas and not others. A light source is positioned opposite to a photocell with the disc between. The photocell is divided into many tracks depending on the resolution required. A typical number of tracks is ten since this would allow 2^{10} possible combinations which would consist of ten photocells (one for each track). The ten bit output would then be amplified and connected to the computer input. Any given temperature would then be represented by a binary number between 0000000000 and 1111111111 which would correspond to 0 to 1023 in decimal terminology.

The actual temperature in engineering units such as degrees F can be determined by the computer through a scaling operation in which the binary number read by the computer is multiplied by a factor determined by the range of the transducer. This factor is determined only once and placed in computer memory. Each reading then is corrected and stored in memory as a temperature value in degrees F.

A clock is included in the system to act as a reference for the sequence of operations in changing temperatures. Data from the clock is continually available to the computer. The clock is a simple counter counting the cycles from the power line since this is quite stable over long periods of time and in any case is as accurate as any synchronous clock. For each sixty counts received, one count is recorded and the first section reset and repeats. The recorded counts are then seconds (at a line frequency of 60 cps). The seconds counting register resets on 60 and transfers a minute count to the next section. Likewise the last section counts hours to 24 and resets. In this way, data is in the proper form to connect directly to the computer and furnish time of day information.

The computer can now be programmed to operate all of the subsystems previously described. A typical program would have the following routines:

1. Bring temperature of oven to first test temperature.
2. Measure capacitances, store and log data.
3. Bring temperature of oven to second test temperature.
4. Measure capacitances, store and log data. (This sequence would continue to the last temperature.)
5. Calculate temperature coefficients and log data.
6. Calculate temperature coefficient distribution and log data.

In order to control the temperature of the oven the computer begins by comparing the desired temperature to the actual. The difference represents an error which is entered into the control equation. Solution of the control equation provides a decision to do one of two things: Turn on power or turn off power to the oven. Clock data also enters at this time and provides a basis for decision as to the stability of the temperature. This is accomplished by comparing the programmed temperature to the actual measured temperature and establishing the time for which the difference remained between ± 2 degrees of zero for example. When this condition has been met for five consecutive minutes, the computer recognizes the fact and initiates a cycle of capacitance measurements.

The capacitance measurement is also made on a cyclic basis from control commands generated by the computer. After the first decision that temperature balance has been sustained for the given length of time, the computer issues a command signal to step the commutator to position 1, then after a small delay (on the order of 10 milliseconds) it commands the frequency counter to take a reading. The frequency counter gate opens for the gate time and then closes. The closing of the gate signifies to the computer that the reading has been completed.

The computer then reads and stores the data from the counter. After calculating the capacitance from the frequency data, the computer then logs out the point number, temperature, and capacitance.

After having completed the routine for the first point, the computer commands the commutator to step to the next point and measure. Again, the same sequence of events would take place until all of the points had been measured, calculated, and logged. The last point is stored in memory and after each capacitance measuring cycle, the computer runs a comparison check for equality between the point just completed and the last point. When the two check as equal, the computer automatically moves to the next temperature.

After a check which indicates that the maximum desired temperature has been reached and the data read and logged, the computer moves into the final routine to calculate the temperature coefficient point by point and log each coefficient along with the point. After completing the coefficient calculations, the program shifts into a statistical analysis of the results and logs out the desired statistical data. Sensing that all data gathering and calculations were complete, the computer would then shut down the test.

The foregoing illustrates the power of a computer system. The system described would not be a large complex of equipment. Figure 6 shows a typical small scale computer which can easily perform the tasks described. It would in fact be possible to package the whole system in a standard relay rack. Since the computer acts rapidly, it can do many chores while waiting for other equipment to perform. For example, in the system described, the computer can make at least one and perhaps many temperature control calculations on the oven temperature while the frequency counter is measuring frequency. The "gate closed" signal from the counter would enter an interrupt input to the computer to indicate that data was ready for pickup by the computer. At the next convenient point in the program, the interrupt would be checked. If it has received a signal, the program would shift into that portion written to operate in the data, and having completed that operation, would return to the point at which it had been interrupted. In this way, the computer is continually juggling all of the inputs and output to gain the maximum amount of computation per unit of time.

Much larger computer systems are presently in operation controlling complete refineries, power generating systems, and cement plants. Such systems not only provide complete automatic operation of entire plant and complexes, but have the added ability to assimilate marketing and other pertinent data such as weather or availability of raw materials data. With this data, the large computer system makes economic decisions and schedule alterations which automatically adjust set point and controls within the plant to case the most economic production of the most marketable item. Versatility of this magnitude stretches the imagination to new horizons in the realm of the possible.

4. A Blending System

In the ceramics industry, the accuracy of material concentrations or mixtures can be quite important to the quality of the end product. Often materials are mixed in a slurry to gain even composition before further processing of the material. The addition of small amounts of reactants with a high degree of accuracy can be a difficult task with conventional means. The following blending system could find use in mixing a reactant solution with a slurry at unusually high accuracies.

Figure 7 shows a typical example of such a system using a variety of control devices. Two feed lines are indicated "A" and "B". "A" is shown to have a magnetic flowmeter since it will be assumed that the material flowing through this flowmeter is abrasive and not suitable for measurement by either positive displacement or turbine meters. "B" is assumed to be a liquid which may be used with a turbine meter. The liquid "B" is to be added to the material "A" in a given ratio by volume which may be set manually.

This system is divided into both analog and digital parts. Since the heart of the control is a pulse counter, the first objective will be to convert the volumetric flows into pulse trains which represent the volume which has passed. In line "A" this is accomplished through the use of a recorder associated with the magnetic flowmeter. Figure 8 pictures a typical magnetic flowmeter. The recorder is usually considered to be a part of the flowmeter due to the sensitive nature of this method of flow measurement, and will indicate the flow rate of the material passing through. An attachment to the recorder will convert the rate reading to a resistance proportional to the rate indicated. A voltage source across the variable resistance will then generate an output voltage proportional to the flow rate. This in turn when connected to a voltage to frequency converter will provide a pulse train whose frequency is proportional to flow rate. Since the frequency is proportional to flow rate, then each pulse must represent a given amount of column. Thus, we have now represented the volume by a series of pulses.

The turbine meter such as that illustrated in Figure 9, generates pulses directly. This is accomplished by placing a small turbine directly in the path of the flowing liquid. As the blades of the turbine pass a reference point on the circumference, a pulse is generated for each blade that passes. Calibration curves furnished with the meter relate flow rate to pulse rate. This can then be interpreted as pulses per unit volume. Since turbine meters are subject to wear both on the blades and in the bearings, the calibration curve is not valid indefinitely. For this reason, periodic calibration checks are run to determine the accuracy of the meter.

When used in a system, one relies upon a single pulse representing a given volume of liquid, therefore a drift in calibration causes loss of accuracy in the control system dependent upon the meter. To compensate for meter wear, electronic scalers have been developed. The scaler is a counting instrument which will generate a preset number of output pulses for every 10^N input pulses, where N is the number of decade units in the scaler.

As an example of how this works, let us examine the case in which a turbine meter normally generates 1000 pulses per gallon. Further, we will assume that a three decade scaler is in use. Since we can never get more pulses out that we put in, we must choose a convenient number of pulses per unit volume with which to work. In this case, 100 pulses per gallon seems to be the most convenient. To obtain this, we must set 100 on the scaler. Since we have three decades, we will get $100/10^N$ or 100/1000 pulses out. Expressed another way, we will get 100 scaler output pulses for every 1000 turbine meter pulses. As wear causes a drop in the number of pulses per unit volume for the flowmeter, we can manually select a larger number (anywhere from 000 to 999) for example 110 which would then cause the generation of more output pulses per 1000 input pulses, but now the meter would be generating fewer pulses per gallon. The result would be that the system would see no change in the pulse per unit volume ratio, and we have thus compensated for the meter wear.

Now data is available from both meter units in the form of pulse trains representing the volume of each that is passing. To use this data, the units must first be normalized to a common language, that is, each pulse must represent the same volume. This can be done by adjusting the voltage span in the voltage to frequency converter to correspond to the chosen pulse-to-volume ratio of the normalized turbine meter.

At this point, it is necessary to discuss the bi-directional counter, an example of which appears in Figure 10, which compares the pulse trains by counting one train in a positive direction while simultaneously counting the other in a negative direction. The result is an instantaneous algebraic sum of all counts received. If both lines "A" and "B" have identical flow rates and both normalized pulse train outputs are fed to the BDC (bi-directional counter), then the constant algebraic sum will be zero, but if one of the lines has an increased flow, the resultant difference will be integrated.

Attached to the BDC is a digital to analog converter which generates an output signal voltage proportional to the number and signal contained in the BDC. The D/A output voltage in turn enters a controller with the set point set to zero. The controller in this instance has proportional plus reset action. Neglecting the Ratio Set control for the moment, consider that the valve in line "B" is closed, and a slurry is pumped through "A". A pulse train would result immediately which would cause the BDC to count away from zero thus causing an output from the D/A converter. The result is that the controller now sees an error signal and acts upon it by opening the valve in line "B". As liquid passes through "B", pulses are generated which tend to reduce the difference count in the BDC and thereby reduce the error signal to the controller. Eventually, the reset action will cause the valve to assume a position such that both flows are equal. This, however, will be after the integrated difference in the BDC has been made up by allowing more of "B" to pass in order to "catch up".

It can be seen that the control system will act upon the valve in the "B" line so as to cause a net count difference in the BDC of zero. For balance, this means that when the system is in operation, both inputs to the BDC must have the same frequency. If it is decided to blend in the ratio of two parts "A" to one part "B", there would need to be a means to compensate the pulse trains to make it appear to the BDC that the pulse rates were equal when the desired ratio existed.

To accomplish ratio control, another scaler is placed in the system and is designated Ratio Set. The Ratio Set control may be present to any value between 00.0% and 99.9%. As an example, assume the setting to be 10.0% which is interpreted to mean that for every gallon of "A" passing through, 0.1 gallons of "B" will be added to the stream. With a setting of 10.0%, there would be an output of 100 pulses for each 1000 input pulses. This in turn means that the command pulse rate to the BDC is 1/10 the pulse rate of the "A" line meter. To reach balance then, the valve will assume a position such that the "B" line meter pulse rate is equal to the Ratio Set output, or 1/10 the rate of the "A"line. This means then that the flow rate of "B" is 10.0% that of "A".

A blending system of the type described lends a greater measure of control to the blending process than other methods due to the integrating action of the bi-directional counter which generates the control action. Since this counter can resolve to a single count, and generate an output signal for that single count, then it is possible to control to within one count representing quite a small volume. The precision in this application is due to the fact that the control action is taking place on volume information and not rate.

5. A Temperature Control System

In a tunnel kiln, the temperature gradient is controlled by zones of different temperatures rather than by changing the temperature as a function of time. For the tunnel kiln, the material is passed from zone to zone as a function of time while the temperature of each zone is held constant.

This process will contain four zones of temperature control. Each zone will have a gas valve or a saturable core reactor to control the amount of heat applied and a thermocouple to measure the temperature. This unit is illustrated in Figure 11. Here is shown the four parts of the control system, the sensor or thermocouple, the signal conditioner or transmitter, the amplifier or controller, and the actuator which may be a valve position or signal level to a saturable core reactor.

The thermocouple is a simple temperature sensor. Its low cost and small size overcome some of its weaknesses such as: low signal level (a few millivolts), need for a cold reference junction (ice bath is ideal but not practical), limited temperature range (depends upon the type of thermocouple used), and the need to place the thermocouple junction at the spot where the temperature is desired.

The signal conditioner or transmitter depends upon the type of sensor being used and the type of system which follows. For the thermocouple example, the transmitter must accept a millivolt input signal, add to this a cold reference junction millivolt signal, and produce the needed output signal. For the case of an electronic analog controller, the output signal would be 1 - 5 milliamperes or some other common 1 to 5 ratio current levels. The output signal at this stage represents the span of the system. For example, if the input is 0 to 5 millivolts after the reference junction representing 0 to 100°C and the ouptut is 1 to 5 milliamperes, the span is set at 0 to 100° Centigrade.

The amplifier or controller is the heart and the brains of the system. An amplifier by itself would be used in only the simplest of systems. The controller contains an amplifier, but also contains a set point signal which is compared with the signal from the transmitter. The signal from the comparison or error signal is amplified and used to produce the signal from the controller. More control action can be obtained by integrating the error signal and/or by taking the derivative of the error signal and adding the result of each to produce the final output. This output for an electronic analog controller would be a 1 to 5 current ratio signal such as 1 to 5 ma.

The 1 to 5 milliampere signal could be used to control the firing angle of a saturable reactor directly, or it could be used to position a valve by possibly going through an electric to pneumatic relay, before driving a pneumatic valve positioner.

The same process control could be accomplished pneumatically.

6. A Direct Digital Control Approach

The four loops of the tunnel kiln indicate the possibility of using a common control system which is time shared among the four loops. All direct digital control systems are designed in this manner in order to minimize the cost per loop of the digital equipment. An example of this type of controller is shown in Figure 12.

Figure 13 shows in block diagram form the operation of a direct digital controller controlling the four temperature loops of Figure 9. Note that the thermocouple temperature sensors and the gas valve actuators are still used. In this case, the direct digital controller takes the place of the transmitters and the analog controllers.

The DDC starts its cycle by selecting the proper input and output scanners. Closing the input scanner connects the thermocouple signal to the temperature reference compensator and to the amplifier. The gain of the amplifier sets the span of the system, the same as the transmitter in the analog system. The next step is to set the desired digital set point into the bi-directional counter. The D/A converter following the bi-directional counter changes the digital number to an analog voltage. This known analog voltage is compared with the unknown voltage from the amplifier by the voltage comparator. The output of the voltage comparator indicates which voltage is larger and thus the direction that the bi-directional counter must be stepped in order to make the known analog voltage equal to the unknown voltage. The direction of stepping and the number of steps are the important factors in this step. The direction represents the polarity and the number of steps, the magnitude of the error in digital form. The number in the bi-directional counter, after completion of a balance, is the digitized form of the temperature. This number is displayed in percent of span of the input millivolt signal. While not directly readable in engineering units, this method is quite sufficient to provide excellent control action.

The output scanner comes into play during the actual analog to digital conversion by connecting the error signal to the proper output circuit. The error signal for this output is converted to + and - voltages representing a positive error or a negative error. Again, the numbers of pulses represent the magnitude of the error. The output circuit integrates these error pulses and either increases or decreases the output current in order to reduce this error to zero. The integrator has a gain adjustment in order to speed up or slow down the control action for each individual loop. The output current is 1 to 5 milliamperes or other standard 1 to 5 ratio current values.

Finally, to complete one cycle of the DDC, the loop number is displayed and the variable under control remains constant until another cycle. This is done by obtaining the loop number from the loop scanner and the variable from the bi-directional counter.

The other three loops are repeats of this cycle, changing only the input and output scanner connections along with different set points. The simplest method of storing a digital set point is to use 3 selector switches for each point. Each switch contains a number in binary coded decimal form by having four decks wired with this code. Another method is to store the set point in twelve flip-flops. A big advantage of the flip-flops is their ability to accept set points from a remote location or from a digital computer.

The direct digital controller has as one of its greatest advantages, the ability to communicate with a digital computer. The A/D converter obtains the variable in digital form which is the form needed by the computer. The computer can either present set points to the controller or can be made to present digital error information in the form of digital pulses directly to the output circuits. This combination becomes a very powerful control tool due to the fast speed and large memory capabilities of the digital computer and the interface, as well as the control functions of the direct digital controller.

Another feature of this approach is the ease with which written data may be obtained even without the use of a computer. Since the set points, variables, and control loop designations are electrically available in a coded form, page copy of the data is conveniently possible. This can be significant in the temperature control of a kiln, especially where gradients are critical to performance and a permanent record of this data would be helpful to the operator.

7. A Computer Controlled System

While the computer has left an indelible mark in the data tabulation field, it has not yet reached its peak in the process control area. Predictions of completely automated plants by visionaries of the last decade are just now appearing, but if present interest in such applications are any indication of their importance, the impact of computers on the field of industrial process and plant control could equal that of the data processing revolution just now at its zenith.

The ceramic industry is using computer technology to control inventory of parts, predict sales and types of product which should be stocked, and even to schedule production runs of products. Recently, however, continuous processes have gone over to complete computer control. Figure 14 illustrates the appearance of a typical computer control console for use in such systems. The first areas within the ceramics field to automate in this respect have been the cement plants. Two processes constitute the major portion of cement manufacture, these are the blending of raw materials, and the kiln-firing of the mixture. Variations exist in the methods of blending materials, however the kiln firing techniques seldom vary.

First, let us consider control methods used in the proportioning of materials. In the case that raw materials are stored in separate hoppers, the control of blend consists of controlling the feeder conveyors. Continuous belt weighing devices can present data to the computer to calculate feed rate, and store the integral in memory in order to account for the total amount entering the process. Feed rates are automatically controlled to correspond to those rates necessary to maintain a mixture ratio of materials set by laboratory analysis of the process output. While this particular approach has the drawback of requiring a laboratory sample analysis to close the loop, new on-line analyzers should allow continuous feedback on future systems.

A second method of proportioning makes use of the control of the stockpile conveyor in order to build a stockpile with a constant mixture cross section which may be crushed and homogenized at the same time. The fact that this proportioning method is off-line allows a more accurate control since adjustment is possible by the addition of materials until the proper balance is reached. The primary drawback of this method is that rapid changes in the product are difficult to accomplish. The scheduling capability of the computer in this case may be able to justify its use, however, it is doubtful that continuous on-line control with analyzer feedback would benefit this method.

The kiln firing should next be considered. As mentioned before, this step in the process seldom varies, however the computerized cement plant may use alternative methods to control the kiln. Two philosophies dominate this area of control technology. Figure 15 illustrates in block diagram form these two methods. Direct computer control is illustrated in 15A which includes various inputs entering the analog to digital converter and from thence into the computer. The digitized information consists of thermocouple signals, fuel flow rate information in the form of standard current signals from transmitters, and perhaps contact closures indicative of various states in the process. In direct computer control, the computer actually performs rapid calculations on the data received from the inputs and relates these calculations to the mathematical model of the kiln. Any adjustments called for as a result of the analysis are generated as output signals in the form of contact closures, pulses to stepper motor drive systems, or time duration signals all of which may be used to adjust valve positions or speed controls.

In this approach, plant shutdown could occur if the computer were to become inoperative. It would be virtually impossible for an operator to operate the control commands normally generated by the computer in an emergency situation. The best that could be hoped for would be that all output commands would remain at the last command which would allow continued operation until the computer could be brought back online. This of course presupposes that the plant can operate satisfactorily on fixed output commands.

The computer directed approach as outlined in 15B, shows that the control is performed in local loops under the direction of controllers and the controller set points are commanded by the computer. This approach permits autonomous local loop control during an emergency loss of computer direction in much the same way that present loop control can be accomplished manually in an emergency condition. The fact that each loop has its own control system means that computer time is saved since it is not necessary to calculate the control equation. Typically, set point adjustment commands are generated as computer output pulses which step a retransmitting slide-wire mechanism. The voltage output of the slidewire represents the desired set point which is digitized and fed back into the computer for verification. Other all-electronic methods are available which eliminate the need for stepper motors and slidewires. Figure 16 illustrates such a typical Output Control Module.

Strong arguments can be made in support of each approach, however, the specific system problem to be solved will usually point to the most useful answer to the problem. A few comments on computer control in general would be useful in grasping the concept of this technology, particularly concerning the common areas of the two philosophies discussed above. Some of the prime reasons for seeking computer control of a plant are the possibility of optimizing plant output, performing data logging, and performing alarm scanning.

Optimization is currently accomplished in two ways. First, and most obvious is the purely mathematical approach in which the mathematical model of the plant is entered, and a program generated which will optimize the control functions based upon the analysis performed in the modeled system. This is undoubtedly the most accurate method, however, as has been mentioned in other papers, the preparation of a mathematical model is the most formidable task in the solution of the optimization program. The second approach, while empirical is more readily programmed. The basis of this approach is the generation of small perturbations which in themselves do not disturb the process significantly, yet result in a process response which will either increase or decrease the plant output. Small set point changes are then made in a direction to increase or optimize plant output. Local loops are randomly selected and adjusted for optimum. While basically simple, this optimizing method has the inherent weakness that the plant might assume maximum output on a combination of control set points which would not represent the most economical input conditions.

Data logging is a natural by-product of computer control. Since all variables and set points are digitized and usually stored in memory, they are all available to a typewriter output. This enables the operator to quickly observe process parameters and at the same time tapes can be prepared to furnish input data to the central tabulating department for accounting purposes. Logging is usually done on a programmed basis, perhaps hourly, unless some emergency situation dictates a logging routine more often.

Alarm scanning is similarly a by-product of computer control. This function normally requires a significant amount of hardware, but with computer control, a simple routine to compare the variable value to the limits stored in memory is all that is necessary to accomplish this task. The change of limits can easily be accomplished at the control console by merely selecting the loop and desired high or low limit, then commanding an entry of the data. The result is that the memory location serving that function assumes a new value, and the comparisons continue searching for deviations. At the time a condition is sensed in which a particular loop exceeded its limits, an alarm is actuated and normally the loop number, time, and data will be logged usually on a separate typewriter. Additionally, a priority interrupt input to the computer will be actuated to cause the computer to quickly attend to this particular situation. The priority interrupt will cause the computer to interrupt its normal program at some convenient point, and switch into an emergency program which will then examine the cause of alarm and initiate corrective action.

While large scale computer control of plants is now in its infancy, it nevertheless has received considerable attention from control users. The results to be seen within the next decade will most likely exceed the strides made in the last by an order of magnitude, and can be expected to greatly affect the lives of each of us.

References

1. "Handbook of Automation, Computation, and Control", Vols. 1, 2, and 3 Grabbe, Ramo, & Wooldridge Editors - Wiley, 1959

2. "Principles of Industrial Process Control", D. P. Eckman, March, 1945

3. "Standard Instrumentation Questions and Answers", Vols. 1 & 2. Steve Elonka and Alonzo R. Parsons, McGraw-Hill, 1962

4. "Instrument for Measurement and Control", Werner G. Holzbock, Reinhold, 1962

5. "100 Process Control Systems", M. H. Aronson and F. D. Marton - Instruments Publishing Company, 1962

6. "Advanced Instrumentation and the Control Computer", H. H. Flum and D. J. Fraade, Thompson Ramo Wooldridge Inc. ISA Preprint Number 22.2.63

7. "Instrumentation for Digital Computer Control of a Cement Plant", Dr. D. D. Bedworth and J. R. Faillace, General Electric Industry Control Department, ISA Preprint Number 22.4.63.

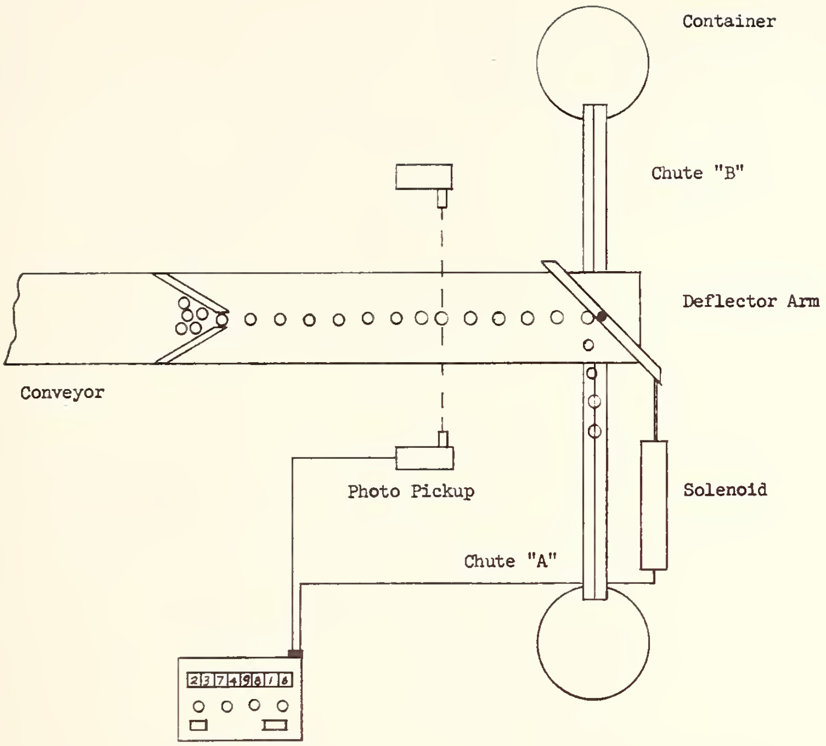


Figure 1. Parts Packaging System

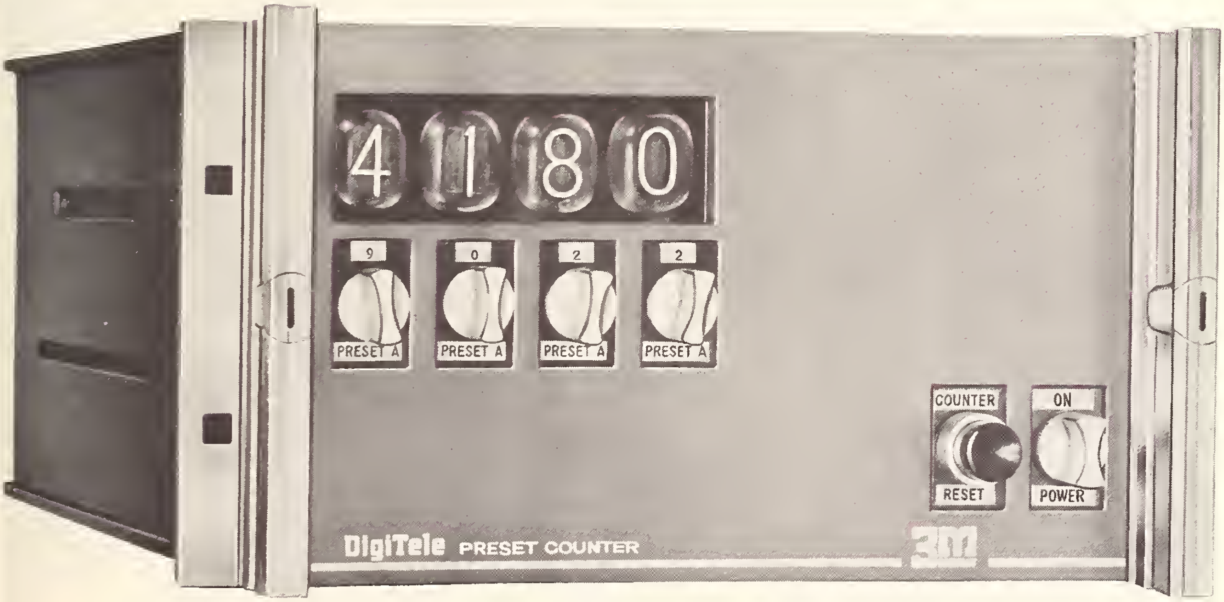


Figure 2. Preset Counter

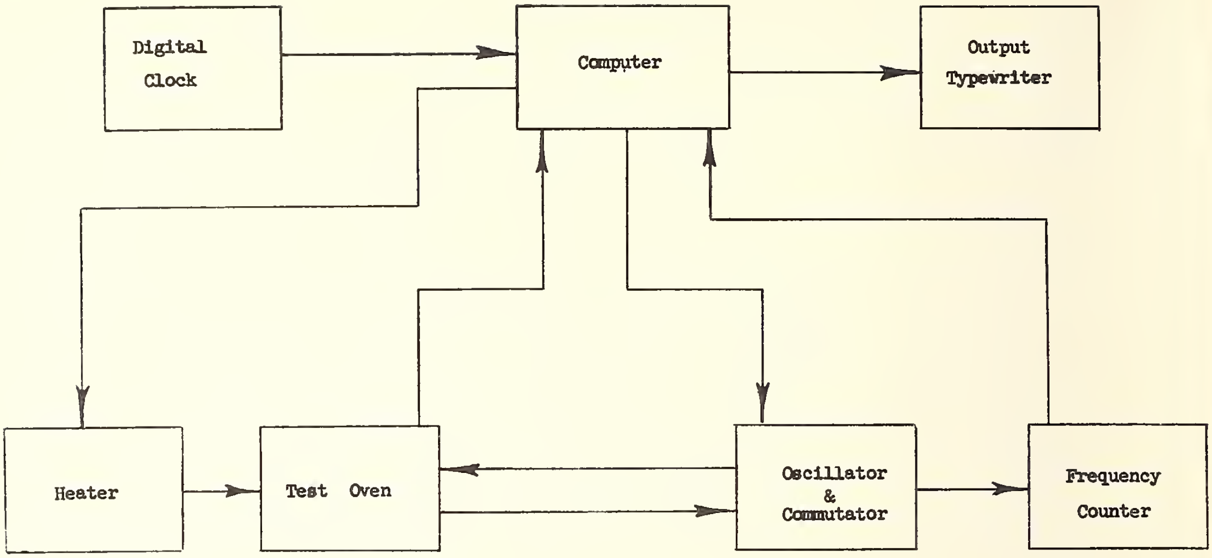


Figure 3. Automatic Testing System

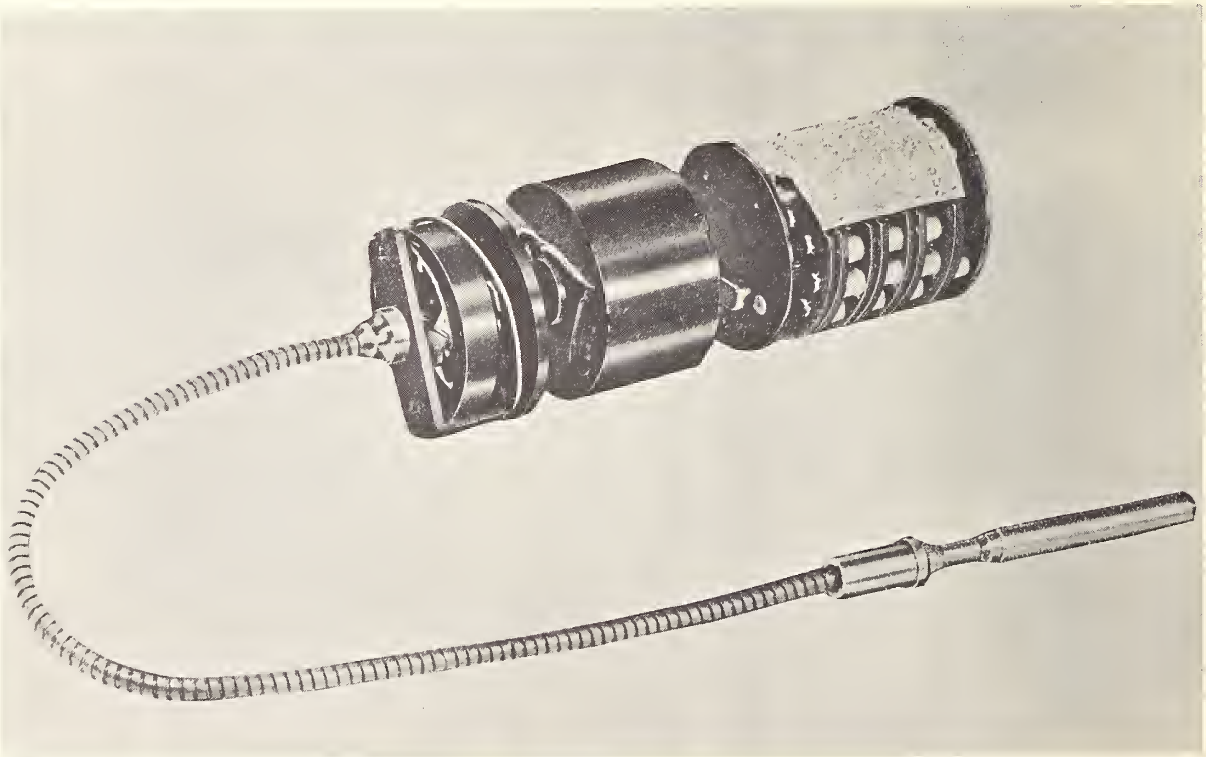


Figure 4. Temperature Transducer/Encoder

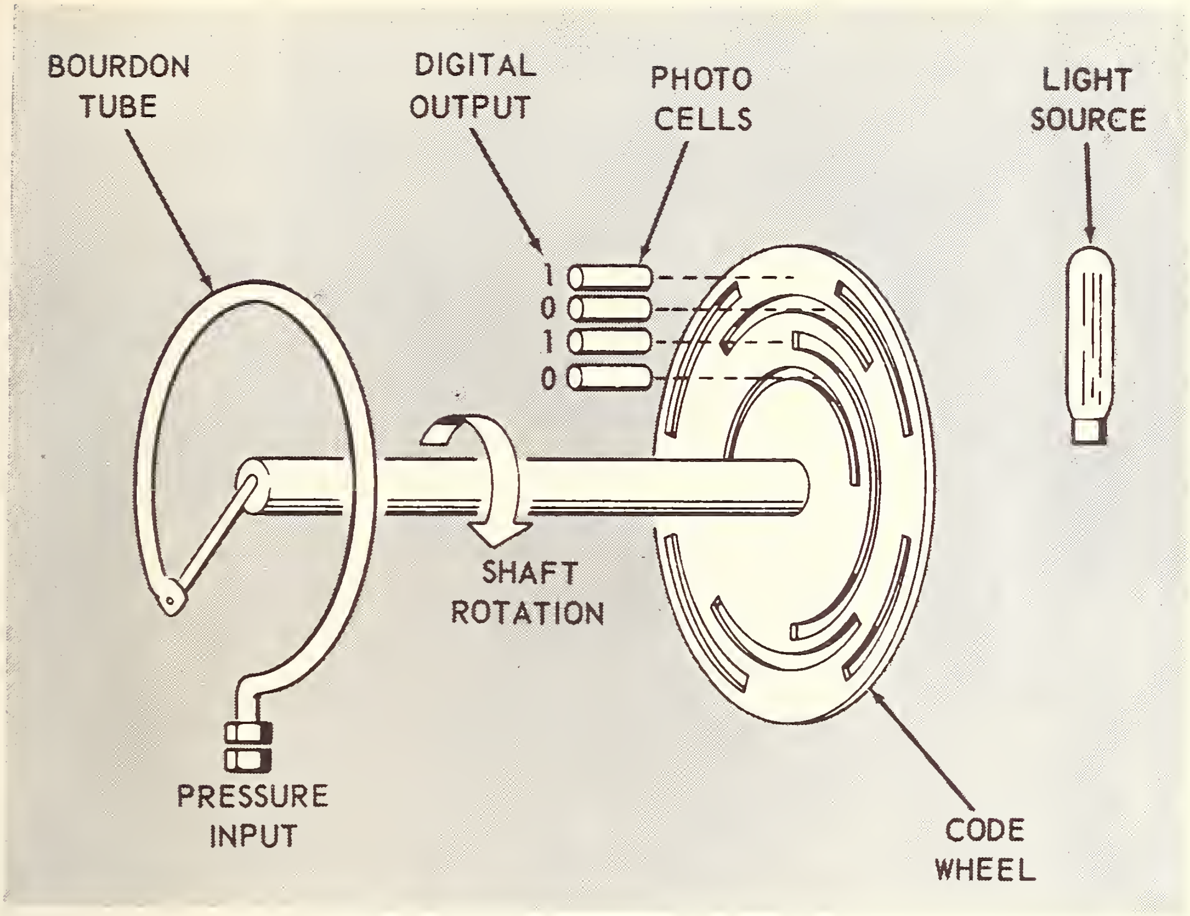


Figure 5. Transducer/Encoder Mechanism

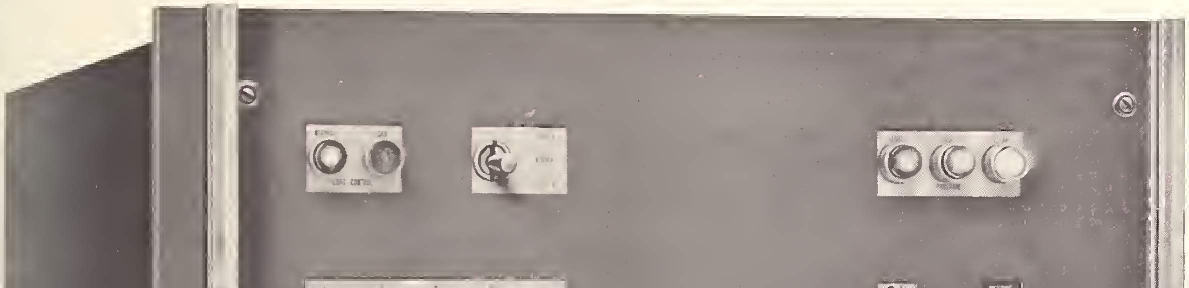


Figure 6. Computer

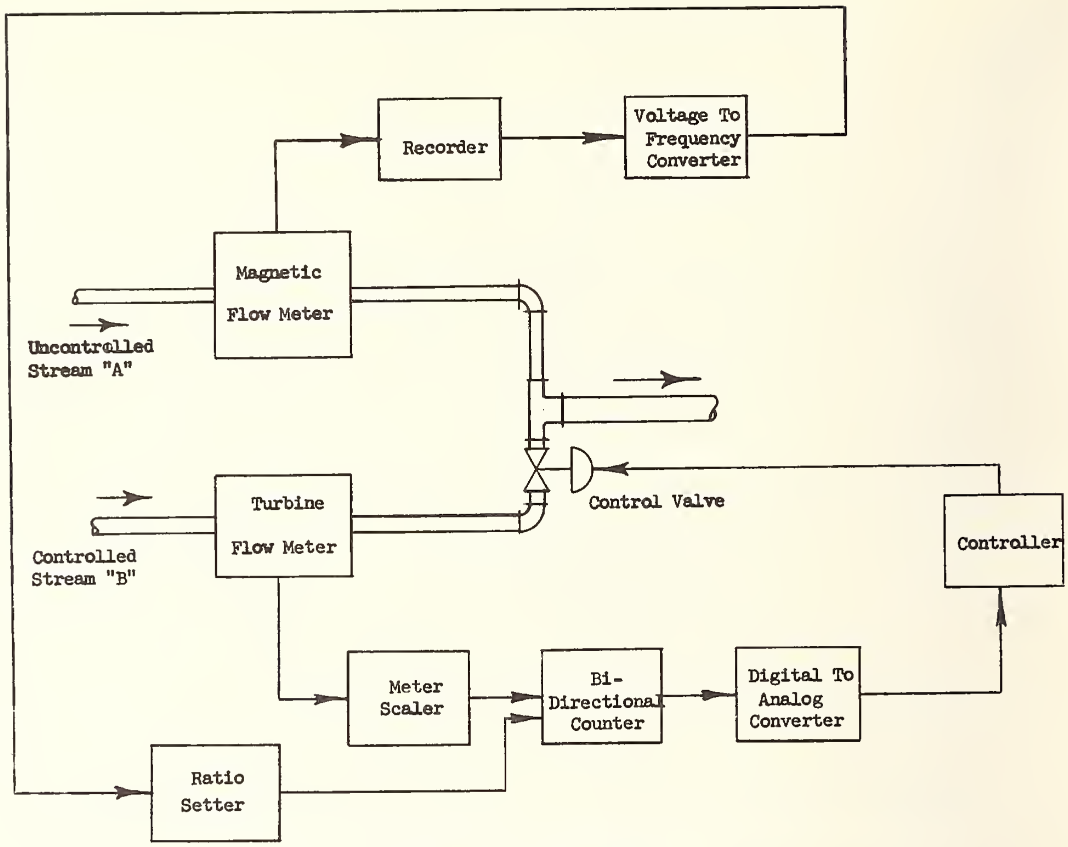


Figure 7. Blending

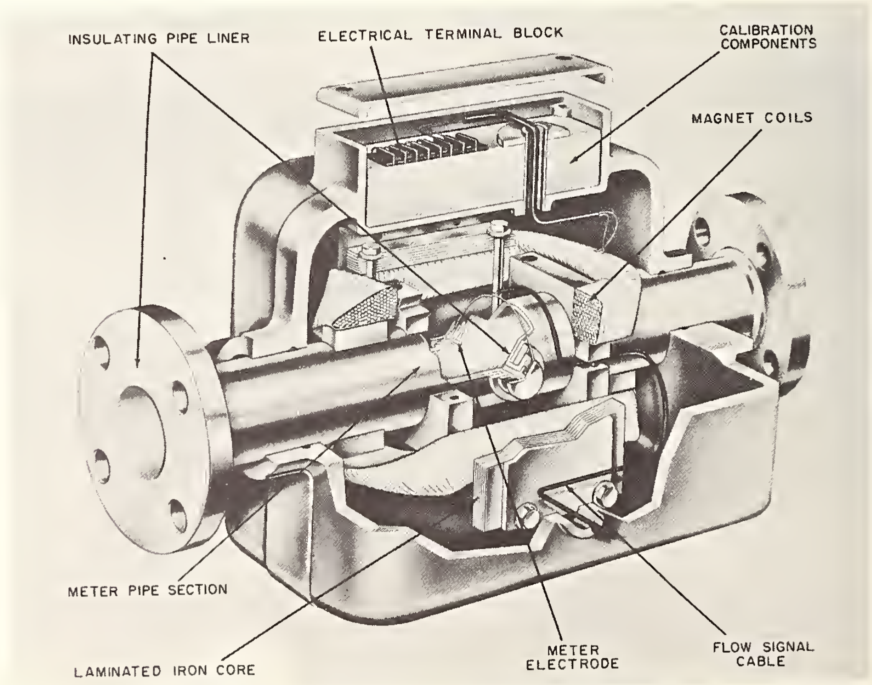


Figure 8. Magnetic Flow Meter (Courtesy Fischer & Porter Company)

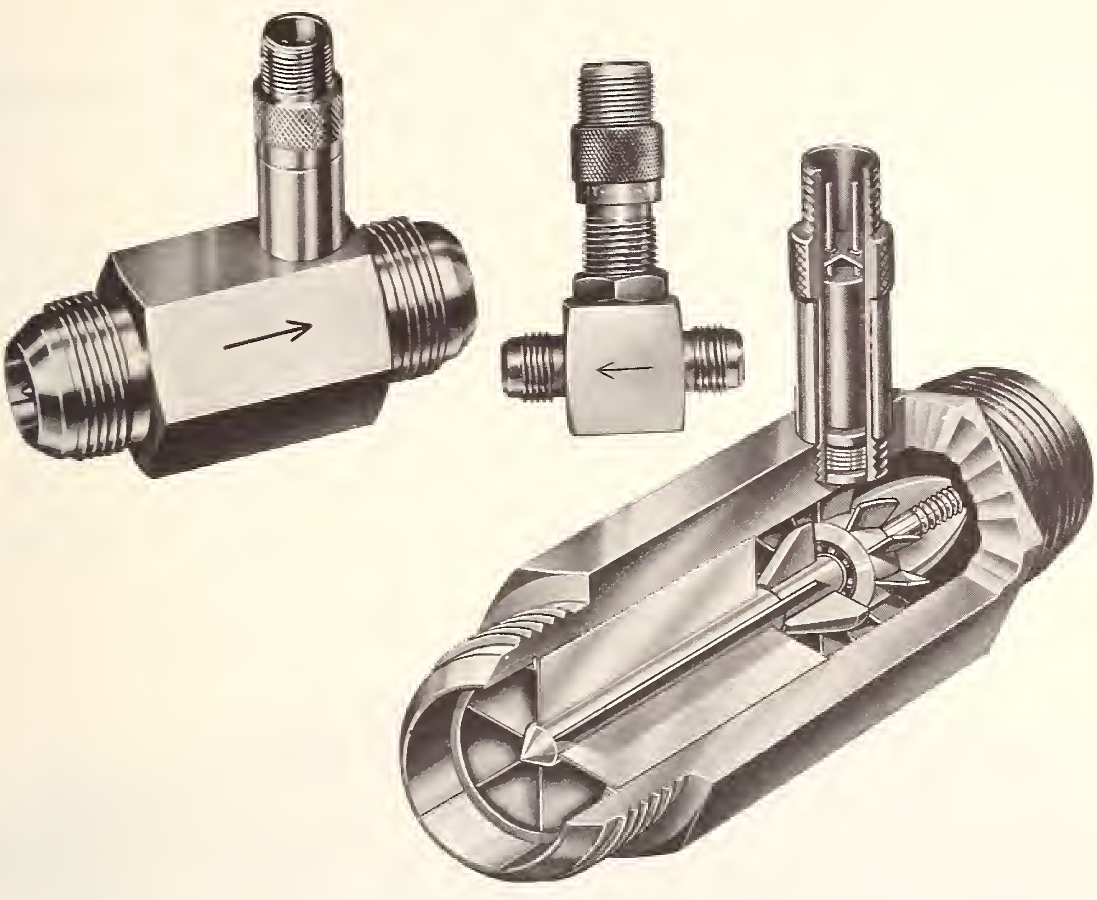


Figure 9. Turbine Meter (Courtesy Cox Instruments Div., Nankervis Co.)

Figure 10. Bi-Directional Counter

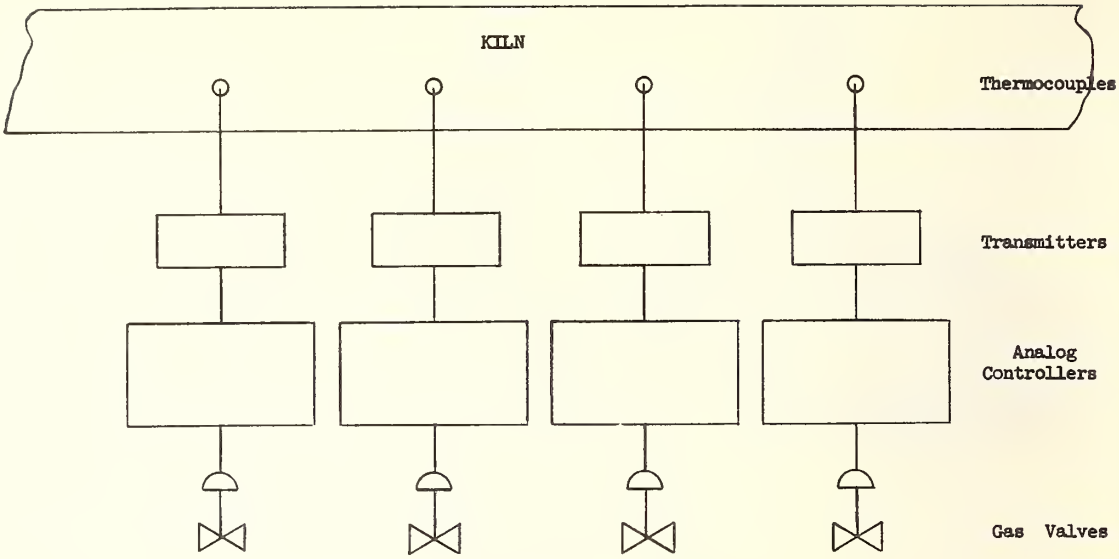


Figure 11. Temperature Control System

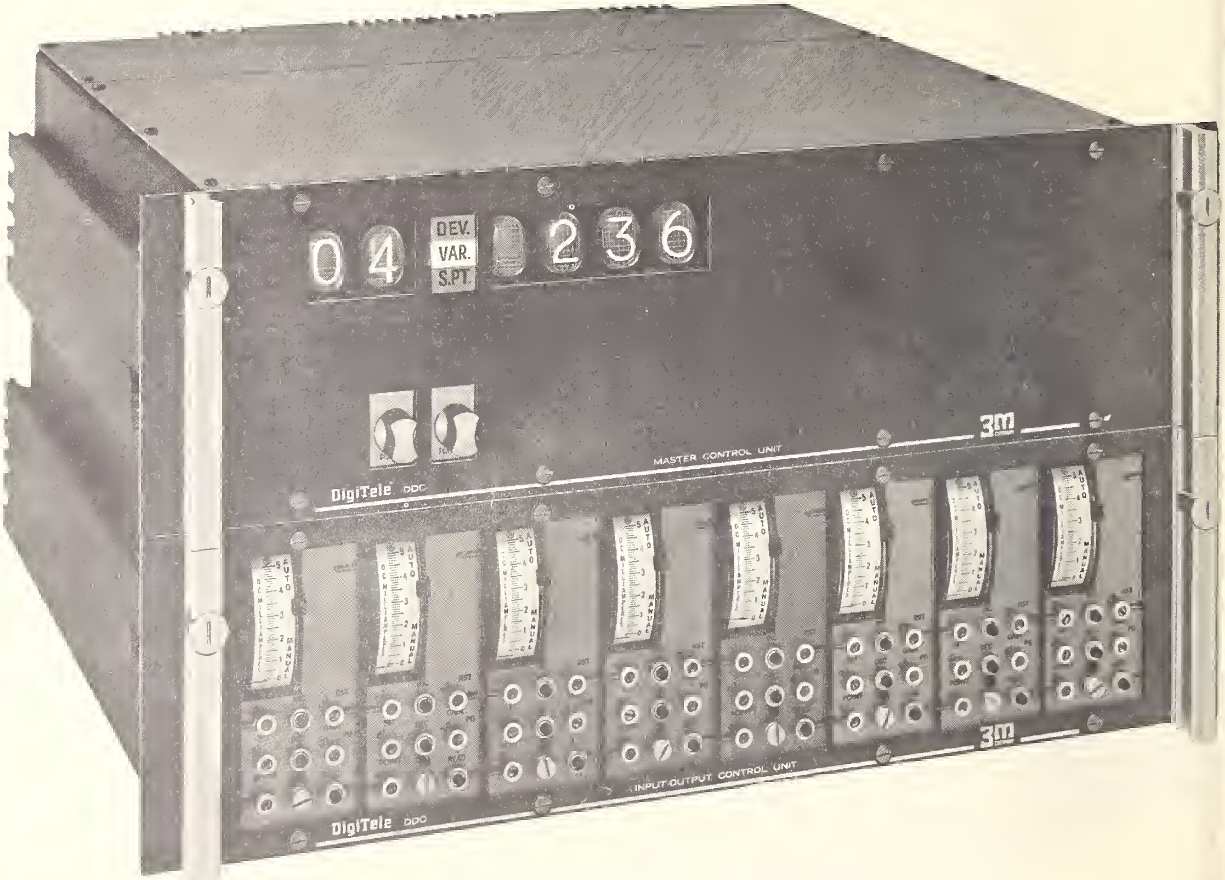


Figure 12. Direct Digital Controller

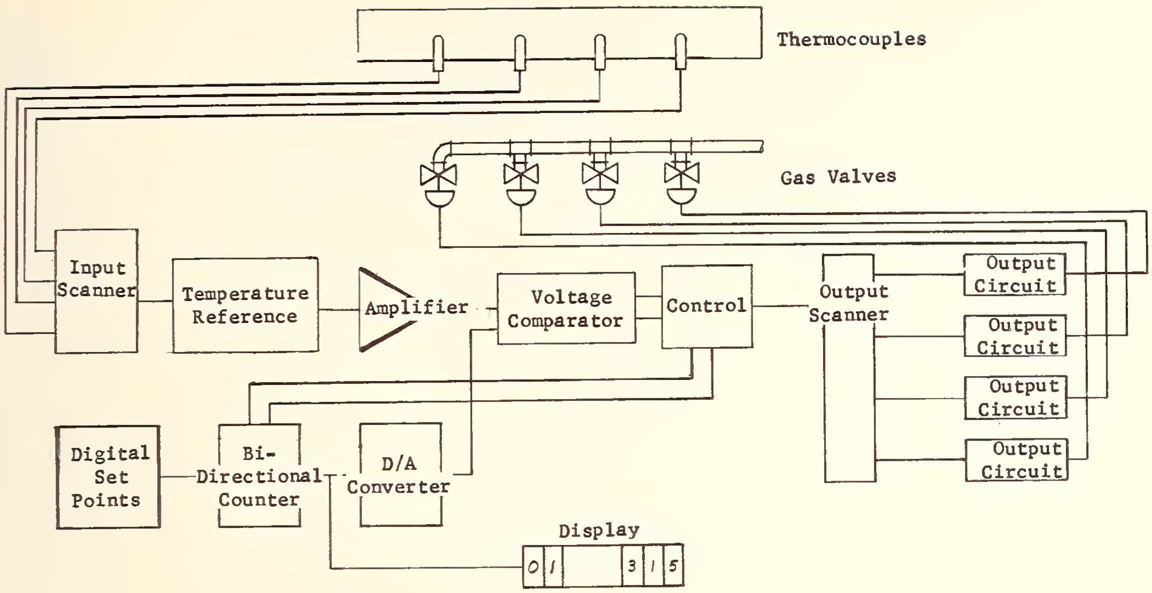


Figure 13. Temperature Control System with Digital Control

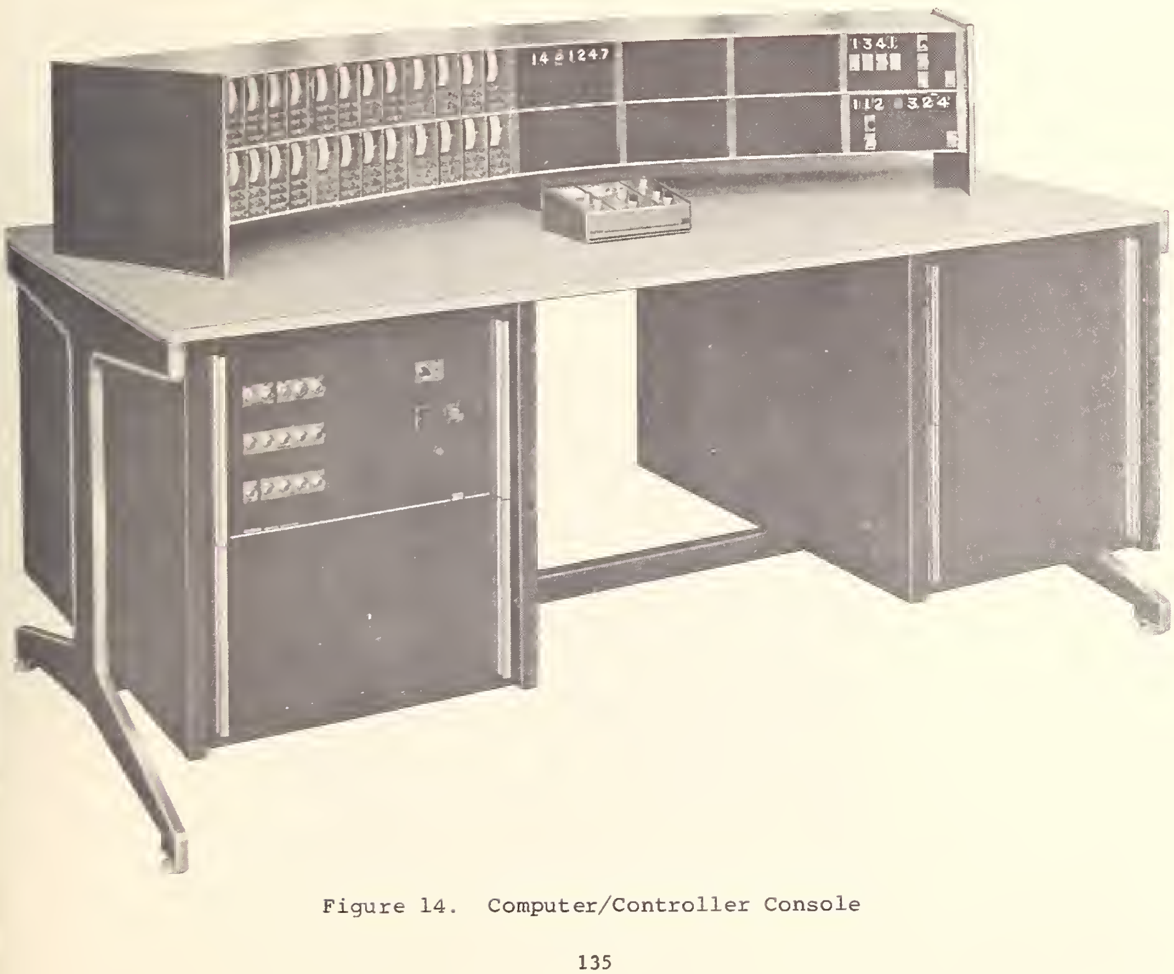


Figure 14. Computer/Controller Console

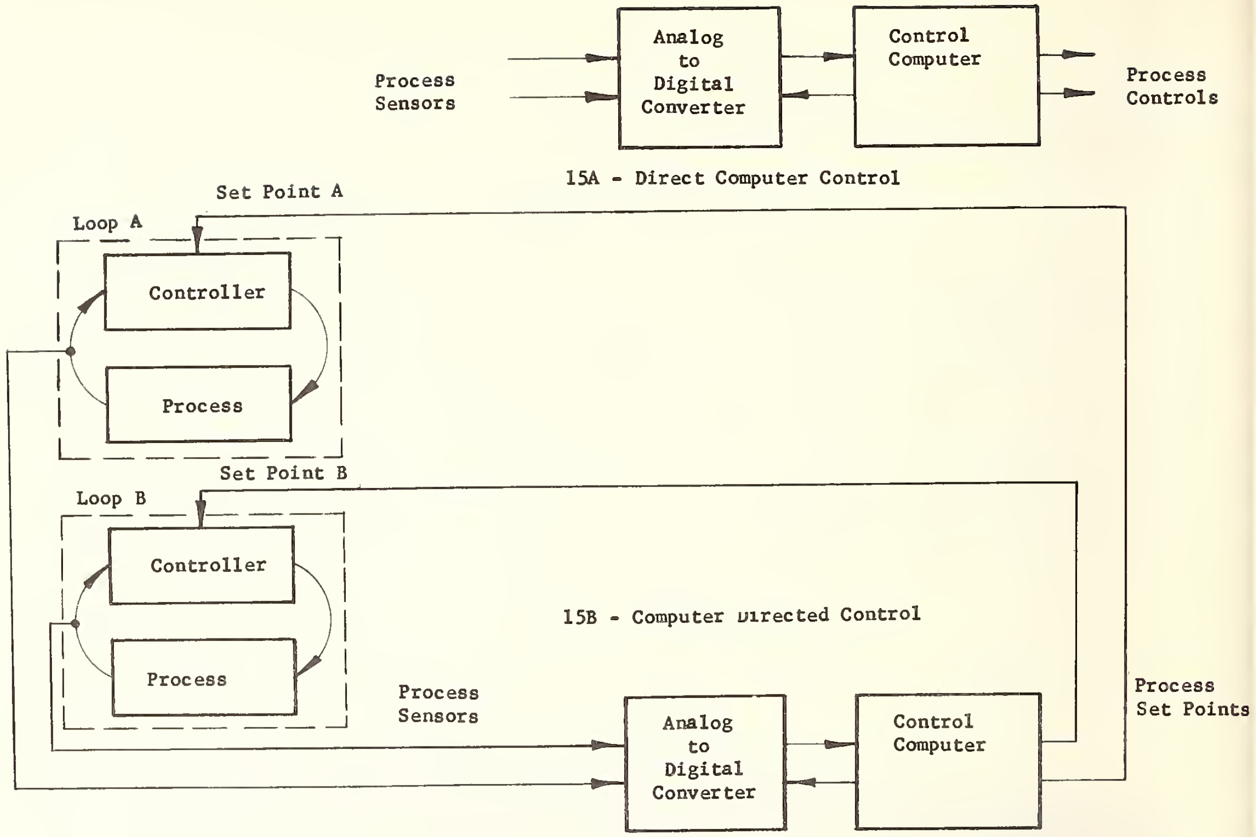


Figure 15. Computer Control System

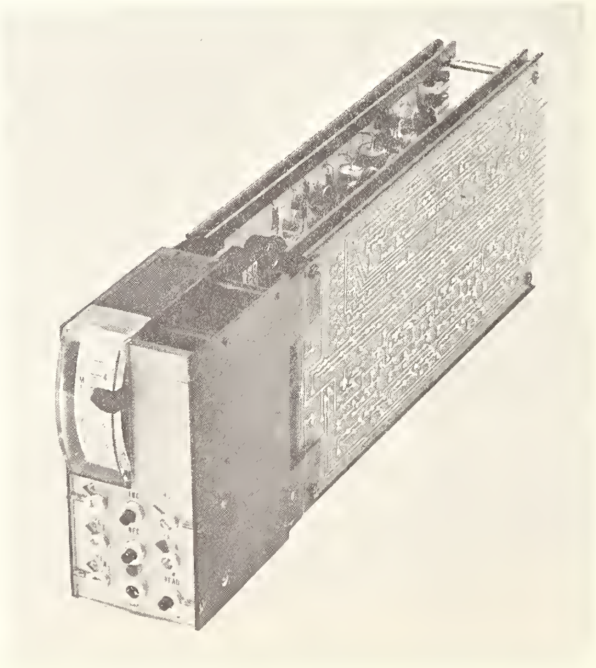


Figure 16. Typical Output Control Module

☆ U. S. GOVERNMENT PRINTING OFFICE : 1965 O - 765-415