Analysis and design of engineering systems : class notes for M.I.T. course 2.751 / by Henry M. Paynter ; with the assistance of Peter Briggs.

Paynter, Henry Martyn, 1923-
Cambridge, Mass. : M.I.T. Press, c1961.

http://hdl.handle.net/2027/mdp.39015064874921

Paynter

Analysis and Design of Engineering Sys

[redacted]

Analysis and Design
of
Engineering Systems

Class Notes for M.I.T. Course 2.751

by

Henry M. Paynter

Associate Professor of Mechanical Engineering
Massachusetts Institute of Technology

The M.I.T. Press
Cambridge, Massachusetts

Analysis and Design *of* Engineering Systems

Class Notes for M.I.T. Course 2.751

by

Henry M. Paynter
Associate Professor of Mechanical Engineering
Massachusetts Institute of Technology

with the assistance of

Peter Briggs

The M.I.T. Press
Cambridge, Massachusetts

Library of Congress Catalog Card Number:

PRINTED IN THE UNITED STATES OF AMERICA

DEDICATION

TO BERN DIBNER in appreciation of his keen sense of scientific history and magnificent standards of aesthetic quality;

TO GEORGE A. PHILBRICK in small token of more than a decade of friendship and service as foil and mentor par excellence;

TO HAROLD A. WHEELER as yet another scholar and craftsman whose seminal concepts were a prime catalyst for this effort.

PREFACE

These notes represent a partial evolution from a primeval jungle of hand-duplicated class notes used over the last seven years in a highly successful one-semester course for seniors and graduate students at the Massachusetts Institute of Technology. More recently, a second semester has been added; thus the reader should be forewarned that the rather abrupt and preclimactic closure of these notes coincides with the end of the <u>first</u> term (2.751) syllabus.

The present complete course is intended to provide both a practical and a theoretical background for engineers of all pedigree who find themselves dealing with complex systems which operate simultaneously in several media and over a broad band of frequency.

The systems here of principal concern are real and material; as such, they may be described in energetic terms. System behavior is so viewed from the standpoint of energy continuity and power balance: a generalized Poynting vector is defined, valid both for continuous and for reticular systems. The useful and significant subsystems are then classified both with respect to the number of energy ports through which energy is exchanged with the environment and also in terms of the particular <u>internal power transformations</u> involved. Thereby the synthesis and design of systems involves a selection and interconnection of a set of standard multiport elements, appropriate to accomplish the required tasks.

It will be noted that both active and passive systems are treated; thus automatic feedback control, together with power and signal amplifiers, are assumed as natural means of providing such activation. Moreover, while energy and power remain central throughout this treatment, signal flow in real devices is considered consistently as a form of low-power-level communication bond. Thus such recondite topics as channel-capacity, gain-bandwidth and indeterminacy, may be treated in a workmanlike fashion.

Electrical engineers should be particularly tolerant in criticism of this approach to engineering systems; but lest they be tempted to condemn this material as mere generalized (or perhaps, worse, *debased*) circuit theor they should remark the central role played here by the philosophical concept of *relations*, (as opposed to functions), by the logical concept of *serial order* (as opposed to measure), and by the engineering concept of *essential multiports* (as opposed to 1-port impedances). Such topics and others beside while largely absent and perhaps superfluous in linear circuit theory, becam foundation stones upon which a general framework of practical theory for engineering systems may be laid with confidence.

Certain of the material presented in these notes will also appear in a forthcoming book *Ergs and Bits: The Flow of Energy and Signals in Engineering Systems* to be published by the McGraw-Hill Book Company.

One should bear in mind that for the M.I.T. students this abbreviated text is richly supplemented with about eighty home and examination problems, covering a broad range of applications, and serving to amplify and to clarify many points only briefly mentioned here. Regrettably, it has proved impractical to include the problem material herewithin.

The reader may thank the diligence of Mr. Peter Briggs for much of the early part of the text; he put many hours into a valiant effort to capture on paper the flavor of previously undocumented lectures. Other particular thanks for comments and corrections are due to many former students and present colleagues, and especially to Mr. David R. Vaughan. Also appreciated was the continuing advice and counsel of Prof. S. A. Coons, Prof. J. L. Shearer, and Prof. T. B. Sheridan.

The manuscript is the work of Mrs. Amy Botelho, with the help of Mrs. Addison Dahmen and Miss. Bonny Davis. The symbols and figures are largely the effort of Miss. Alice Griffin of the M. I. T. Illustration Service. Of course, all faults and blemishes must be laid to the author who assures the reader that criticism will be respectfully welcomed.

Henry M. Paynter

Cambridge, Massachusetts
June, 1961

TABLE OF CONTENTS

Part I. Introductory Remarks 1
- A. Engineering Systems 1
- B. Engineering and Pure Science 2
- C. Specific Examples of Engineering Systems 3
- D. Systems and Abstractions 10

Part II. The System Concept: Identity, Structure, and Properties 11
- A. Description of a System 11
- B. Reticulation 13

Part III. Variables and Parameters of Energetic Systems 17
- A. Introduction and Historical Background 17
- B. Description of Energy Transactions 17

Part IV. The Continuity of Energy 26
- A. Introduction and Historical Background 26
- B. Properties of the Field 29
- C. Generalized Continuity Equation 30
- D. Continuity of Energy and the Generalized Poynting Vector 31
- E. Continuity of Entropy 32

Part V. Energy Ports and Power Bonds 35
- A. Introduction 35
- B. Application of the Divergence Theorem 36
- C. Reticulation of the Energy Influx 36
- D. Reticulation of Energy Storage 37
- E. Reticulation of Energy Dissipation 38
- F. The Reticulated Equation of Energy Continuity 39
- G. Power Bonds 42

Part VI. Multiported Systems and Elements 49
- A. Introduction 49
- B. Multiports 50
- C. Ideal Energy Junctions 51

Part VII. Classes and Relations 58
- A. Relations and Structure 58
- B. The Concept of a Class 59
- C. The Concept of a Relation 64

Part VIII. Continuum Logic and Hyperpolyhedral Functions 77
A. Introduction 77
B. Classes 77
C. Order 81
D. Continuum Logic 84
E. Two-Valued or Binary Logic 87
F. Multivalued Logic (Post Logics) 91
G. Hyperpolyhedral Functions 92

Part IX. The Steady-State of Energetic Systems 100
A. Introduction 100
B. The Static Case 100
C. The Stationary Case 101
D. Determination of the Steady-State 102
E. System Reticulation for Steady-State Behavior 103
F. Nonlinearity 105

Part X. Functional Transformations and Computing Functionals 108
A. Introduction 108
B. Computing Functionals 108

Part XI. Diagrams and the Coding of System Structure 113
A. Signs 113
B. Communication of a Computing Structure 114
C. Combinatorial Topology - Incidence Matrix 117
D. Coded Representation of Graphs and Digraphs 123
E. Coding the Energetic Structure of Multiport Systems 125

Part XII. State-Determined Systems 130
A. Introduction 130
B. Elements of a State-Determined System 132
C. The Mathematical Construct of a State-Determined System 133
D. The Variables of State in Generalized Dynamics 135
E. The Tetrahedron of State 136
F. The Characteristic Static Relations 137
G. The Three State-Determined Elements (R, C, I) 143
H. The Concept of Circuits and Networks 148

Part XIII. Distribution of Energy Over Space, Time, and Frequency 153
A. Introduction 153
B. Energy Principles for State-Determined Systems 154
C. Fields, Potentials, and Transmittances 159
D. Field Form Factors 178
E. Rectangle Diagrams 186
F. The Temporal Response of Physical Systems 193
G. System Response in the Time and Frequency Domains 198
H. Linear System Response in Terms of Potential Functions 211
I. One-Port Elements and the Impedance Concept 214
J. The Flow of Power and Energy in Systems 219

Part XIV. Two-Port Elements and Energy Transport Processes 223
A. Generalized Two-Port Elements 223
B. Primitive Energy Transport Processes 225
C. Linear Two-Port Elements 227
D. Some Standard Forms of Two-Port Nets 233
E. Description of Linear Two-Ports 235

Part XV. Transformers and Transducers 240
A. The Concept of Ideal Two-Port Elements 240
B. Energy Transformation Elements 241
C. Energy Transduction Elements 250

Part XVI. Energy Transmission Elements 256
A. The Two-Port Element: ·TM· 256
B. The Canonical Transmitter Matrix 259
C. Generalized Transmitters and Wavelike Transmitters 263
D. Ideal Wavelike Transmitters 267
E. Modeling Diffusive Transmission 277
F. The Dynamics of Monotone Processes 280

Part XVII. Energy Modulation and Amplification 287
A. General Three-Port Elements 287
B. Generalized Power Modulators as Three-Port Elements 288
C. Generalized Amplifiers 290
D. The Trinode as a Three-Port Element 293
E. Cascading and Feedback of Amplifiers 299

Analysis and Design

of

Engineering Systems

I. Introductory Remarks

A. Engineering Systems

Characteristics and Classifications

At the outset it may in general be stated that an engineering system is conceived, designed, and constructed to perform a specific task. The content of this course will be concerned with material or physical systems--machines, structures, instruments--which are to be distinguished from the more abstract, nonphysical systems such as economic or social complexes. However, this latter type is no less real than the concrete, physical system; indeed, it is conceivable that the nation's economy might be modeled by delineating the dynamic interaction among elements analogous to the inertial, dissipative, and elastic elements of physical systems.

The following are a few examples of material engineering systems of the type we wish to consider:

1. Services and utilities--water supply, electric power generation, communication;
2. Structures--buildings, houses, bridges;
3. Instruments--clocks, computers;
4. Vehicles--submarines, aircraft, spacecraft, ships, automobiles*.

There are, of course, transcendental systems whose boundaries encompass two or more of the above. A large missile, for example, is necessarily a complex structural system as well as a vehicle, and in addition, it requires electronic computing elements to effect its guidance and control. By the same token, it is often the case that one of the above types may be viewed as an integration of well-defined and physically distinct subsystems; in many vehicles the power plant may be thought of as a propulsion system, conceived apart from the system as a whole.

---

* The latter two are examples of interfacial vehicles.

Engineering systems are historically improved and proliferated. Improvements may occur as a result of lighter, faster, more compact, or more reliable components being substituted for older ones; or as a result of a reconception, redesign, or rearrangement of configuration which utilizes the same or similar components but yields a more perfectly integrated whole. There is, in addition, an all-pervasive trend towards greater sophistication which demands a higher level of ability among those who are responsible for the conception and design of new systems.

B. Engineering and Pure Science

Engineering generally predates science. Historically, invention, the art of the utilization of natural phenomena to realize certain practical objectives, has preceded science which subjects the phenomena to rational analysis and seeks a thorough understanding thereof. Indeed, it has often been the case that an inventor, who has employed incorrect reasoning and analysis, has succeeded in contriving a workable system! The conception and design of the first airplanes was certainly not founded on a thorough understanding of aerodynamic lift. In fact, lift was not fully understood until long after the first successful flight. To the contrary, some early mathematical analyses predicted that airplanes could not fly!

This perhaps dramatizes the fact that traditional modes of analysis often fail to account for those crucial factors which limit the performance of a system, or even those very agencies which enable the system to operate at all. As mentioned, the first analyses of flight overlooked the true nature of the fluid circulation which results in a lift force or sidethrust. More generally, it is the failure to properly and completely account for the flow of matter, energy, information, entropy, etc. which is the downfall of most classical analyses. Biological systems are elusive for this very reason. The well-heralded "second industrial revolution" could conceivably occur once these flow phenomena are more perfectly understood.

Analysis and Design *of* Engineering Systems

Class Notes for M.I.T. Course 2.751

by

Henry M. Paynter

Associate Professor of Mechanical Engineering
Massachusetts Institute of Technology

with the assistance of

Peter Briggs

The M.I.T. Press
Cambridge, Massachusetts

Library of Congress Catalog Card Number:

PRINTED IN THE UNITED STATES OF AMERICA

DEDICATION

TO BERN DIBNER in appreciation of his keen sense of scientific history and magnificent standards of aesthetic quality;

TO GEORGE A. PHILBRICK in small token of more than a decade of friendship and service as foil and mentor par excellence;

TO HAROLD A. WHEELER as yet another scholar and craftsman whose seminal concepts were a prime catalyst for this effort.

PREFACE

These notes represent a partial evolution from a primeval jungle of hand-duplicated class notes used over the last seven years in a highly successful one-semester course for seniors and graduate students at the Massachusetts Institute of Technology. More recently, a second semester has been added; thus the reader should be forewarned that the rather abrupt and preclimactic closure of these notes coincides with the end of the first term (2.751) syllabus.

The present complete course is intended to provide both a practical and a theoretical background for engineers of all pedigree who find themselves dealing with complex systems which operate simultaneously in several media and over a broad band of frequency.

The systems here of principal concern are real and material; as such, they may be described in energetic terms. System behavior is so viewed from the standpoint of energy continuity and power balance: a generalized Poynting vector is defined, valid both for continuous and for reticular systems. The useful and significant subsystems are then classified both with respect to the number of energy ports through which energy is exchanged with the environment and also in terms of the particular internal power transformations involved. Thereby the synthesis and design of systems involves a selection and interconnection of a set of standard multiport elements, appropriate to accomplish the required tasks.

It will be noted that both active and passive systems are treated; thus automatic feedback control, together with power and signal amplifiers, are assumed as natural means of providing such activation. Moreover, while energy and power remain central throughout this treatment, signal flow in real devices is considered consistently as a form of low-power-level communication bond. Thus such recondite topics as channel-capacity, gain-bandwidth and indeterminacy, may be treated in a workmanlike fashion.

Electrical engineers should be particularly tolerant in criticism of this approach to engineering systems; but lest they be tempted to condemn this material as mere generalized (or perhaps, worse, debased) circuit theor they should remark the central role played here by the philosophical concept of relations, (as opposed to functions), by the logical concept of serial order (as opposed to measure), and by the engineering concept of essential multiports (as opposed to 1-port impedances). Such topics and others beside while largely absent and perhaps superfluous in linear circuit theory, becam foundation stones upon which a general framework of practical theory for engineering systems may be laid with confidence.

Certain of the material presented in these notes will also appear in a forthcoming book Ergs and Bits: The Flow of Energy and Signals in Engineering Systems to be published by the McGraw-Hill Book Company.

One should bear in mind that for the M.I.T. students this abbreviated text is richly supplemented with about eighty home and examination problems, covering a broad range of applications, and serving to amplify and to clarify many points only briefly mentioned here. Regrettably, it has proved impractical to include the problem material herewithin.

The reader may thank the diligence of Mr. Peter Briggs for much of the early part of the text; he put many hours into a valiant effort to capture on paper the flavor of previously undocumented lectures. Other particular thanks for comments and corrections are due to many former students and present colleagues, and especially to Mr. David R. Vaughan. Also appreciated was the continuing advice and counsel of Prof. S. A. Coons, Prof. J. L. Shearer, and Prof. T. B. Sheridan.

The manuscript is the work of Mrs. Amy Botelho, with the help of Mrs. Addison Dahmen and Miss. Bonny Davis. The symbols and figures are largely the effort of Miss. Alice Griffin of the M. I. T. Illustration Service. Of course, all faults and blemishes must be laid to the author who assures the reader that criticism will be respectfully welcomed.

Henry M. Paynter

Cambridge, Massachusetts
June, 1961

TABLE OF CONTENTS

Part I. Introductory Remarks 1
- A. Engineering Systems 1
- B. Engineering and Pure Science 2
- C. Specific Examples of Engineering Systems 3
- D. Systems and Abstractions 10

Part II. The System Concept: Identity, Structure, and Properties 11
- A. Description of a System 11
- B. Reticulation 13

Part III. Variables and Parameters of Energetic Systems 17
- A. Introduction and Historical Background 17
- B. Description of Energy Transactions 17

Part IV. The Continuity of Energy 26
- A. Introduction and Historical Background 26
- B. Properties of the Field 29
- C. Generalized Continuity Equation 30
- D. Continuity of Energy and the Generalized Poynting Vector 31
- E. Continuity of Entropy 32

Part V. Energy Ports and Power Bonds 35
- A. Introduction 35
- B. Application of the Divergence Theorem 36
- C. Reticulation of the Energy Influx 36
- D. Reticulation of Energy Storage 37
- E. Reticulation of Energy Dissipation 38
- F. The Reticulated Equation of Energy Continuity 39
- G. Power Bonds 42

Part VI. Multiported Systems and Elements 49
- A. Introduction 49
- B. Multiports 50
- C. Ideal Energy Junctions 51

Part VII. Classes and Relations 58
- A. Relations and Structure 58
- B. The Concept of a Class 59
- C. The Concept of a Relation 64

Part VIII. Continuum Logic and Hyperpolyhedral Functions 77
A. Introduction 77
B. Classes 77
C. Order 81
D. Continuum Logic 84
E. Two-Valued or Binary Logic 87
F. Multivalued Logic (Post Logics) 91
G. Hyperpolyhedral Functions 92

Part IX. The Steady-State of Energetic Systems 100
A. Introduction 100
B. The Static Case 100
C. The Stationary Case 101
D. Determination of the Steady-State 102
E. System Reticulation for Steady-State Behavior 103
F. Nonlinearity 105

Part X. Functional Transformations and Computing Functionals 108
A. Introduction 108
B. Computing Functionals 108

Part XI. Diagrams and the Coding of System Structure 113
A. Signs 113
B. Communication of a Computing Structure 114
C. Combinatorial Topology - Incidence Matrix 117
D. Coded Representation of Graphs and Digraphs 123
E. Coding the Energetic Structure of Multiport Systems 125

Part XII. State-Determined Systems 130
A. Introduction 130
B. Elements of a State-Determined System 132
C. The Mathematical Construct of a State-Determined System 133
D. The Variables of State in Generalized Dynamics 135
E. The Tetrahedron of State 136
F. The Characteristic Static Relations 137
G. The Three State-Determined Elements (R, C, I) 143
H. The Concept of Circuits and Networks 148

Part XIII. Distribution of Energy Over Space, Time, and Frequency 153
A. Introduction 153
B. Energy Principles for State-Determined Systems 154
C. Fields, Potentials, and Transmittances 159
D. Field Form Factors 178
E. Rectangle Diagrams 186
F. The Temporal Response of Physical Systems 193
G. System Response in the Time and Frequency Domains 198
H. Linear System Response in Terms of Potential Functions 211
I. One-Port Elements and the Impedance Concept 214
J. The Flow of Power and Energy in Systems 219

Part XIV. Two-Port Elements and Energy Transport Processes 223
A. Generalized Two-Port Elements 223
B. Primitive Energy Transport Processes 225
C. Linear Two-Port Elements 227
D. Some Standard Forms of Two-Port Nets 233
E. Description of Linear Two-Ports 235

Part XV. Transformers and Transducers 240
A. The Concept of Ideal Two-Port Elements 240
B. Energy Transformation Elements 241
C. Energy Transduction Elements 250

Part XVI. Energy Transmission Elements 256
A. The Two-Port Element: ·TM· 256
B. The Canonical Transmitter Matrix 259
C. Generalized Transmitters and Wavelike Transmitters 263
D. Ideal Wavelike Transmitters 267
E. Modeling Diffusive Transmission 277
F. The Dynamics of Monotone Processes 280

Part XVII. Energy Modulation and Amplification 287
A. General Three-Port Elements 287
B. Generalized Power Modulators as Three-Port Elements 288
C. Generalized Amplifiers 290
D. The Trinode as a Three-Port Element 293
E. Cascading and Feedback of Amplifiers 299

Analysis and Design

of

Engineering Systems

I. Introductory Remarks

A. Engineering Systems

Characteristics and Classifications

At the outset it may in general be stated that an engineering system is conceived, designed, and constructed to perform a specific task. The content of this course will be concerned with material or physical systems--machines, structures, instruments--which are to be distinguished from the more abstract, nonphysical systems such as economic or social complexes. However, this latter type is no less real than the concrete, physical system; indeed, it is conceivable that the nation's economy might be modeled by delineating the dynamic interaction among elements analogous to the inertial, dissipative, and elastic elements of physical systems.

The following are a few examples of material engineering systems of the type we wish to consider:

1. Services and utilities--water supply, electric power generation, communication;
2. Structures--buildings, houses, bridges;
3. Instruments--clocks, computers;
4. Vehicles--submarines, aircraft, spacecraft, ships, automobiles*.

There are, of course, transcendental systems whose boundaries encompass two or more of the above. A large missile, for example, is necessarily a complex structural system as well as a vehicle, and in addition, it requires electronic computing elements to effect its guidance and control. By the same token, it is often the case that one of the above types may be viewed as an integration of well-defined and physically distinct subsystems; in many vehicles the power plant may be thought of as a propulsion system, conceived apart from the system as a whole.

---

* The latter two are examples of interfacial vehicles.

Engineering systems are historically improved and proliferated. Improvements may occur as a result of lighter, faster, more compact, or more reliable components being substituted for older ones; or as a result of a reconception, redesign, or rearrangement of configuration which utilizes the same or similar components but yields a more perfectly integrated whole. There is, in addition, an all-pervasive trend towards greater sophistication which demands a higher level of ability among those who are responsible for the conception and design of new systems.

B. Engineering and Pure Science

Engineering generally predates science. Historically, invention, the art of the utilization of natural phenomena to realize certain practical objectives, has preceded science which subjects the phenomena to rational analysis and seeks a thorough understanding thereof. Indeed, it has often been the case that an inventor, who has employed incorrect reasoning and analysis, has succeeded in contriving a workable system! The conception and design of the first airplanes was certainly not founded on a thorough understanding of aerodynamic lift. In fact, lift was not fully understood until long after the first successful flight. To the contrary, some early mathematical analyses predicted that airplanes could not fly!

This perhaps dramatizes the fact that traditional modes of analysis often fail to account for those crucial factors which limit the performance of a system, or even those very agencies which enable the system to operate at all. As mentioned, the first analyses of flight overlooked the true nature of the fluid circulation which results in a lift force or sidethrust. More generally, it is the failure to properly and completely account for the flow of matter, energy, information, entropy, etc. which is the downfall of most classical analyses. Biological systems are elusive for this very reason. The well-heralded "second industrial revolution" could conceivably occur once these flow phenomena are more perfectly understood.

C. Specific Examples of Engineering Systems

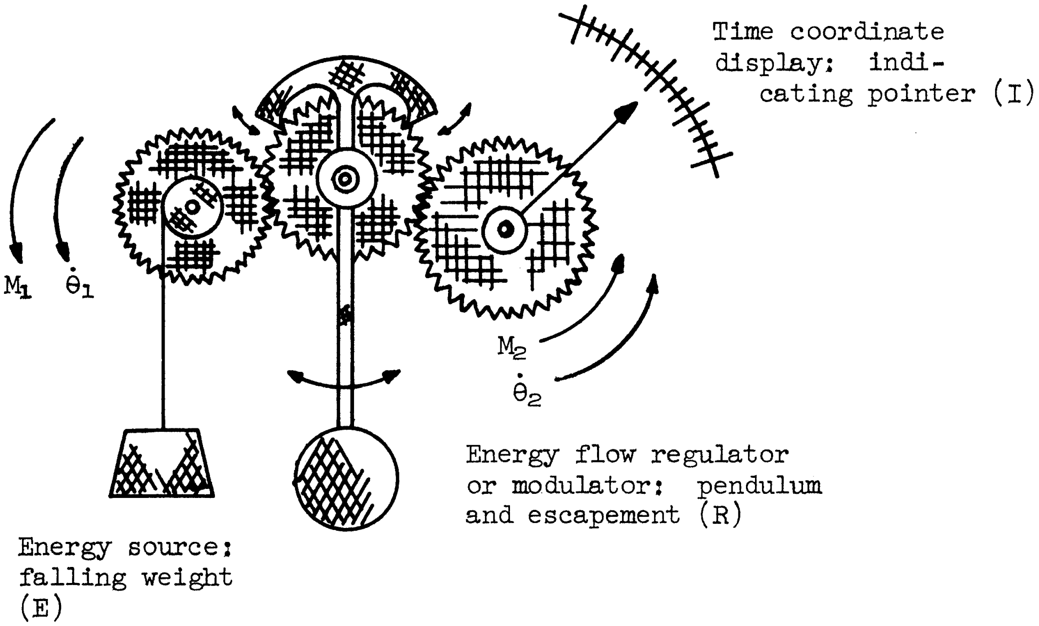
So as to clarify what is meant by the "analysis and design of engineering systems" and to point out the objectives toward which such analysis will be directed, two examples will now be considered.

The Clock

Men have traditionally relied on the celestial clock as the standard time keeper; however, within the past few years we have been able to manufacture clocks which are sufficiently precise to enable a measurement of the inaccuracies and variabilities of celestial time as observed from the noisy and unstable platform of the earth's surface. Thus, the ultimate time keeper will undoubtedly be an engineering system, i.e., an instrument whose sole purpose is to indicate a running time coordinate.

Let us consider the common pendulum clock, reducing it to its basic elements and identifying the variables with which the interactions among the elements may be delineated. It is worthwhile to point out that by removing the veil of material embodiment given to a particular type of clock--the pendulum clock--it will become evident that clocks in general consist of three basic elements: (i) a source of energy; (ii) a gate or regulator of the flow of energy from this source; (iii) an indicator to "read out" or display the desired running time coordinate. The concept "clock" is thus reduced to a schematic diagram showing the flow of energy through its essential elements. Such a schematic simplifies the analysis of the system and facilitates the incorporation of improved elements into the structure as these become available.

Now, the pendulum, by itself, is a nonlinear damped oscillating system; in particular, its period depends upon the amplitude. Thus by merely initiating an oscillation and counting subsequent decrementing swings we cannot hope to achieve the type of time keeper we desire--one that continually reads out a running time coordinate--due to the fact that the free pendulum is a transient device inherently nonlinear and of "irregular periodicity." Moreover, we are, of course, unable to resolve the swings as the amplitude approaches zero. However, by providing a



$$\left\{ E \right\} \frac{M_1(t)}{\dot{\theta}_1(t)} \left\{ R \right\} \frac{M_2(t)}{\dot{\theta}_2(t)} \left\{ I \right\}$$

A PENDULUM CLOCK

Figure 1

source of energy and a gate to regulate the in-flow of this energy, it is possible to maintain the amplitude of the pendulum nearly uniform; then with the incorporation of an indicating pointer intermittently driven by the energy source under control of the oscillating pendulum a true clock is achieved.

This flow of energy through the system is crucial. Thus, we identify the variables $M_1(t)$, $\dot{\theta}_1(t)$, and $M_2(t)$, $\dot{\theta}_2(t)$--the torques and angular velocities at each of the two interfaces within the system. The product $M\dot{\theta}$ is the instantaneous power or time rate of energy flow. Therefore

$$\mathbb{P}_1(t) = M_1(t) \cdot \dot{\theta}_1(t) = M_1(t) \cdot \omega_1(t)$$

$$\mathbb{P}_2(t) = M_2(t) \cdot \dot{\theta}_2(t) = M_2(t) \cdot \omega_2(t)$$

P_1 and P_2 are thus the instantaneous rates of energy flow at the two interfaces. Usually power is conceived as a time-and-space averaged invariant, but we shall see that such an interpretation leads to inconsistencies in any analysis predicated thereon.

Let us scrutinize the E-R interface; it is apparent that the torque $M_1(t)$ is dependent upon the velocity, $\omega_1(t)$. Thus we may regard $\omega_1(t)$ as an "input" to the energy source and $M_1(t)$ as an "output." The roles are necessarily reversed from the viewpoint of the pendulum-escapement. Thus, a more precise delineation of the interaction at this interface might be sketched as follows:

$$\begin{tikzcd}\left\{E\right\} \arrow[r, shift left, "M_1(t)"] & \left\{R\right\} \arrow[l, shift left, "\omega_1(t)"]\end{tikzcd}$$

The ideas introduced in this example will be dealt with further in the second example--an analysis of a typical vehicular propulsion system.

Vehicle Propulsion

The purpose of a vehicle is to transfer itself and its cargo from one point in space to another. The path along which the vehicle may travel in executing a given mission is usually subject to many constraints; these arise, for example, as a result of the necessity to maximize economy, avoid collision with other vehicles or stationary objects, etc. The generalized vehicle problem may be viewed as two component problems, both of which involve the concept of control or regulation. It is first necessary to control the attitude or orientation of the vehicle; secondly, it is necessary to regulate the speed and path or trajectory. The solution to the second problem is embodied in the vehicular propulsion system. We shall now rather cursorily analyze the propulsion system of a ship merely as a further illustration of the techniques to be employed in this course.

Some basic assumptions must first be made. It will be supposed that we are studying a ship whose hull shape and power plant have been fixed. Presumably, these have been chosen so as to achieve optimum economy and performance in light of the service for which the ship is intended. For example, let the power plant be an oil-fired steam boiler with a direct turbo-drive.

A fundamental problem which confronts the propulsion system analyst is this: how does the ship respond to a command to accelerate from one speed to another? The duration and nature of the transient is, needless to say, a function of the power plant and its associated controls, the hull shape and surface preparation, and the state of the fluid in which the ship is floating. Three possible transients are sketched in Figure 2.

The propulsion system of the vessel may be represented schematically as Figure 3-a. Isolating the elements of the system and delineating the interactions among them by way of energy bonds is indicated in Figure 3-b.

Again, it is well to point out that each of these energy bonds refers to a local energy state in the system; for example, the bond between the screw and the reduction gear may be labeled (M,ω) where M is the torque in the drive shaft and ω its angular velocity. Thus, the power flowing <u>to</u> the screw is

$$\mathbb{P}_1 = M \cdot \omega$$

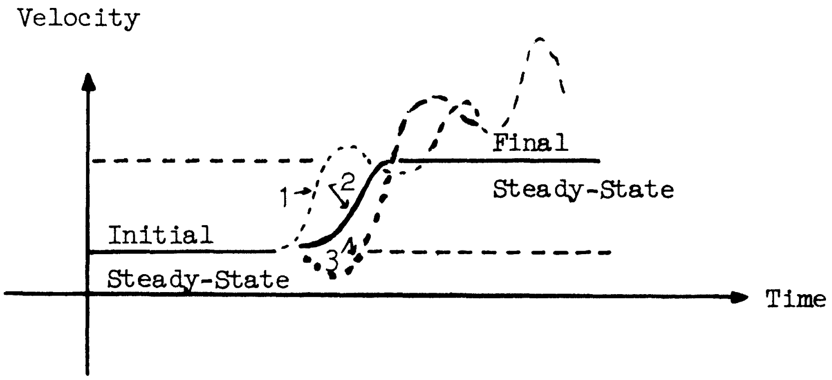

Velocity
Final
Steady-State
1
2
Initial
Steady-State
3
4
Time

Figure 2

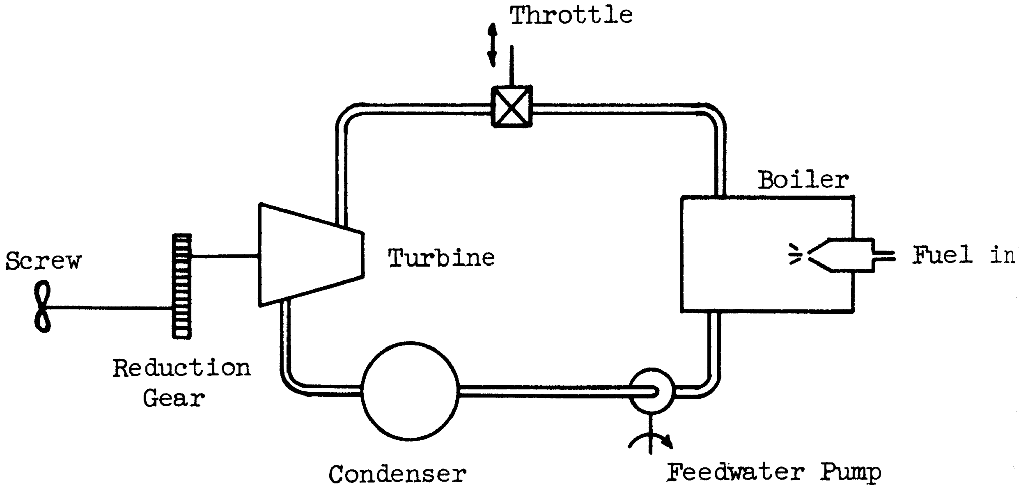


Figure 3-a

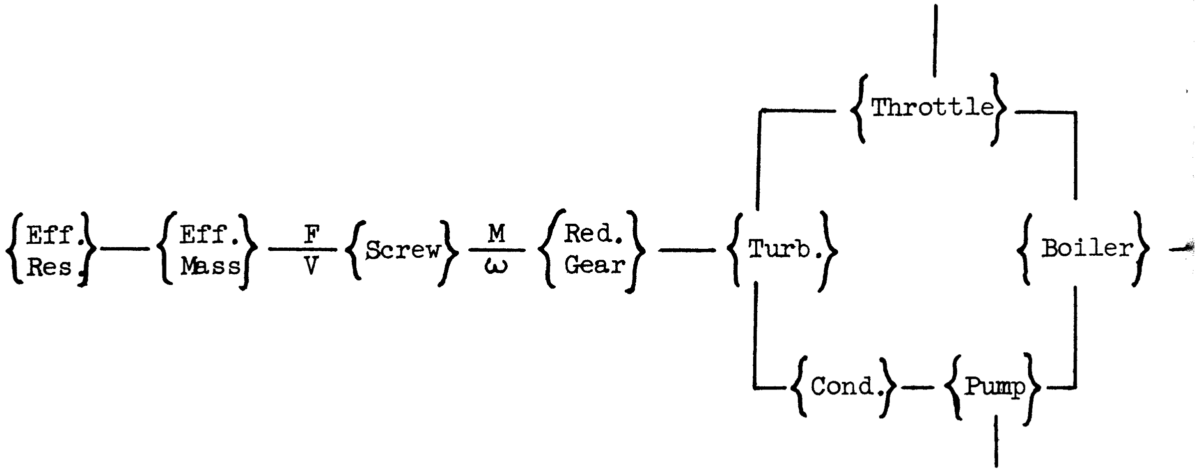


Figure 3-b

A bond is shown between the screw and the <u>effective mass</u> of the ship (the mass of the ship plus the virtual mass of the water moving with the ship); this bond is labeled F, V, so that the power flowing <u>from</u> the screw is

$$\mathbb{P}_o = F \cdot V$$

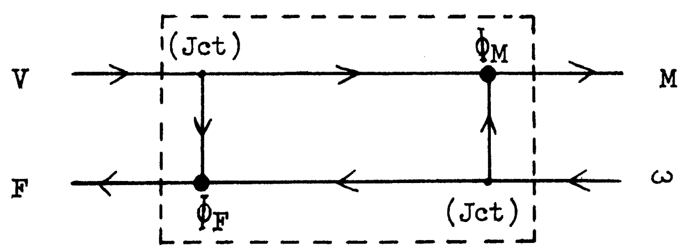
Finally, the available energy imparted to the effective ship mass is ultimately dissipated by way of an <u>effective resistance</u> which accounts for the frictional resistance acting on the hull itself as well as a virtual resistance resulting from eddy formation in the surrounding fluid.

Consider the energy state on either side of the screw driven at constant speed ω while the ship is moving at a constant velocity V. The screw is a transducer or two-port device with $\mathbb{P}_i = M \cdot \omega$ flowing in and $\mathbb{P}_o = F \cdot V$ flowing out. Now, only a certain two of the four variables, F, V, M, ω, may be considered as independent or input variables, while the remaining two are labeled dependent or outputs. Suppose, for example, that we choose V, ω, as the independent variables. Then there must exist two relations:

$$F = \Phi_F(V, \omega)$$
$$M = \Phi_M(V, \omega)$$

which are characteristics of the screw and describe completely its behavior. A causally directed signal flow graph of the screw might then be sketched as follows:

Ship Screw



Such a diagram both delineates the signal flow from and to the screw and also states the constraints upon the variables. However, the same information is embodied in the following simpler diagram:

This tells us that the two "flow" variables V, ω, are "inputs" to the screw while the two "effort" variables, F, M, are "outputs" from the screw and therefore "inputs" to the adjacent elements. The technique of formulating and manipulating such bond diagrams will be a primary concern of this course.

D. Systems and Abstractions

Some degree of abstraction is inevitable in the analysis of any physical system. The second example considered above illustrates this point nicely; in particular, we modeled the effect of the fluid contiguous to the vessel in terms of a virtual mass and a virtual resistance. In so doing the important interactions between the ship and the fluid were isolated and shown as localized energy bonds in the conceptual schematic of the propulsion system.

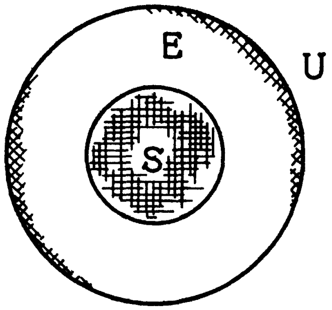
The artful act of abstracting from the totality of interactions between the elements of a physical system and the elements of its environment, and from among the various parts of the system itself, only those interactions which are relevant to the specific questions being asked, and then expressing these mathematically, is certainly a crucial step in the analysis of system behavior. Once this has been accomplished we are no longer talking about the physical system which is the subject of the analysis but rather about a conceived, abstract substitute system or model which is embodied in the mathematical relationships connecting its parts. Thus, an appreciation of the properties of abstract systems is indeed a prerequisite to the incisive analysis of physical systems.

II. The System Concept: Identity, Structure, and Properties

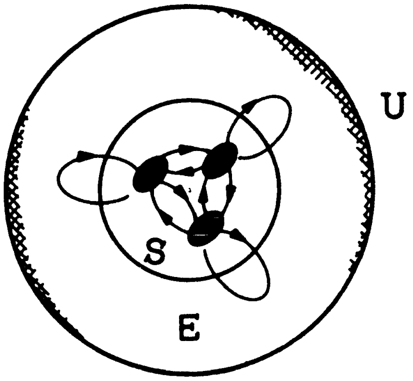
A. Description of a System

Although material systems are of primary interest in this course, the inevitability of abstraction as a step in the analysis of a material system was pointed out at the end of Part I. We have also alluded to systems which are purely abstract. Thus, in formulating an approach to the description of a system, it behooves us to concentrate on the properties and governing relationships possessed by all systems--both material and abstract. Once a sufficiently general systems theory has been developed, the task of specializing to a particular type will be correspondingly simpler.

The description of a system necessarily begins with the identification of a universe (U)--or "universe of discourse" as it is often called in logic--which is a domain or set of sufficient scope to include all elements within the system, plus all exterior elements with which the system may be interacting. The system (S) is then a well-defined subset of U, the elements of S possessing properties and interrelationships which happen to be of particular interest to us. The complementary set, U less S, we shall label the environment E: hence, E is the set of elements which interact with the elements of S but are not in S.



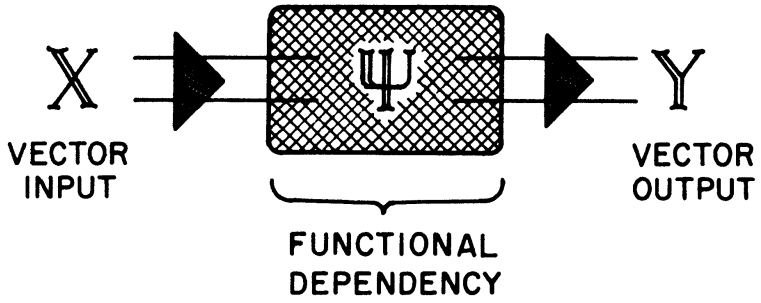
The rational process of endowing a system with structure we call reticulation. The act of separating S from U, and thus defining the interface between S and E, is the first step in this process. The system is further reticulated by conceptually tearing it apart into its essential elements. Since the structural attribute of a system which interests us most is the functional connectedness of its elements, the final step in the reticulation process is the sketching of the important relations and bonds of interaction among the elements and between each of the elements and the environment:



Further description of a properly reticulated system involves a more careful delineation of the functional connectedness of its elements. For the purposes of our general treatment of material systems, this may be accomplished by either of two techniques:

1. An energy bond may be conceived as an interaction; associated with each bond are two variables, the first pertaining to an effort and the second to a flow, their product yielding the power or energy flow rate.
2. Alternatively, an interaction may be conceived as a bilateral signal flow between two elements, thus attributing a direction of causality to the interaction.

A reticulated energy bond diagram facilitates a general understanding of the functional connectedness of a material system and therefore might well be used as a tool in synthesis and preliminary analysis. However, for the purposes of a more detailed analysis the noncausal energy bond reticulation is usually not adequate and must be transformed into the bilateral signal flow reticulation. The description of the system is then completed by conceptually substituting for each element a black box for which the input-output functional dependency is specified.



B. Reticulation

The Latin form, reti, meaning (fish) net, is the stem from which the word reticulation is derived. The verb, to reticulate, means literally to make into or like a network. Hence, for our use, reticulation is a peculiarly vivid term to impart the idea of a conscious act of structuring in the form of a network. How better can a system be characterized than as a group of elements tenaciously bonded to one another as are the meshpoints of a net?

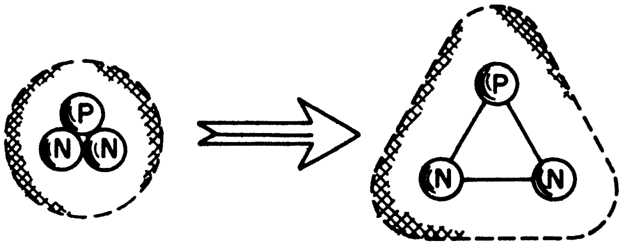
We have said that reticulation begins with the partitioning of the universe (U) into two subsets, the system (S) and the environment (E). It should be emphasized that the environment can only be defined relative to the system; that is, a conjugate relationship exists between E and S such that E is defined once U has been identified and those elements belonging to S have been marked off.

The system communicates with its environment across the E-S interface. Unfortunately, traditional modes of analysis have often placed constraints upon the forms in which such communication can occur. For example, classical thermodynamics deals entirely with systems whose boundaries are permeable only to matter and energy; thus, all living processes, which continuously import negative entropy and information, are perforce excluded from conventional thermodynamical analysis unless the existence of a class of truly open systems is admitted.

The interface between E and S is inevitably a function of the observer and, in particular, the questions he intends to ask concerning the behavior of the system. Generally it is not established by the surface of a material embodiment as might be anticipated in the case of a simple physical system. For example, consider a straight-backed chair. A practical analysis of a chair would certainly not be concerned with the isolated object "chair;" instead, it would consider the interaction of this object with the seat of a human occupant, with the floor on which it rests, etc. Thus the demarcation of the system boundary is far more sophisticated than might superficially appear.

Reticulation has been defined as the process by which a system is endowed with structure. This process is facilitated if we view the system as the integration of a number of subsystems. Thus, the elements of a system are simply systems of lower order, and therefore interactions occurring between two elements are of the same class as those which may occur between two systems. Likewise, it is convenient to view the environment as being structured to the extent that all S-E interactions may be viewed as occurring between two systems--one an element of S and the other of E. Hence, all possible interactions occurring within S or between S and E may, in fact, be looked upon as system-to-system interactions.

We have implied that a system is, in truth, a hierarchy of subsystems. It is interesting to note, in the case of engineering or man-made material systems, that there is a level of decomposition such that all systems of a lower order are natural systems (See Hall and Fagen). In some instances this level is rather low, a case in point being that of a thermonuclear device employing tritium as the fusable material. Since the "element" tritium is itself man-made, the reticulation would have to be carried all the way to the nuclear particles--in particular, a proton and two neutrons. The tritium nucleus is certainly a system in every respect; it may be reticulated into three elements with energy bonds linking these elements.



Although we lack sufficient knowledge to further reticulate a "proton" or a "neutron," there is no a priori reason for denying that such a reticulation could be performed. Hence, we can only say that a further reticulation makes little sense in the context of this course, since at this base level the fundamental properties of the elementary constituents must be <u>assumed</u>.

Background Reading

1. General Systems Theory Yearbook, Vol. I. Introduction by Ludwig von BERTALANFFY.

   The requirements of a "general systems theory" are discussed. Also, profound insights into the nature of open and closed systems are given (see excerpts from pages 3-5).

2. KRON, Gabriel. Tensors for Circuits.

   Read the two Introductions, one by HOFFMAN and the other by KRON himself, for some appreciation of KRON's method of "tearing" (diakoptics).

3. HALL and FAGEN. Definition of a System.

   This reference imparts a general understanding of the ideas dealt with in Part II. Sections 1-10, 11.1, and 13, are particularly valuable.

III. Variables and Parameters of Energetic Systems

A. Introduction and Historical Background

At about the turn of the century the analysis of the macroscopic behavior of systems on an energetic basis had gained considerable stature. However, the advent of quantum and relativistic mechanics deflated the theoretical structure to such a degree that the study of Energetics per se lay dormant until the World War II era, at which time it received renewed attention. The more recent treatments re-evaluated the theory as it had previously stood and, for that matter, criticized the well-entrenched theories of classical thermodynamics. Quoting from the foreword of BRØNSTED's Principles and Problems in Energetics:

> Brønsted's Energetics is not to be confused with the energetics associated with the names of Helm, Mach, and Ostwald of the decade 1890-1900. The original proposals failed for many reasons; at least in part in not recognizing the necessity of the coupling of processes and in some cases an incorrect assignment of the intensity and capacity factors of the energy which still persist in textbooks published during the past year. Brønsted recognized these pitfalls and has been most careful to avoid them.

The analysis of material systems on an energy basis will prove most fruitful and illuminating in the context of the present course.

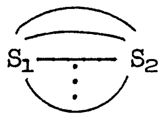
B. Description of Energy Transactions

In the diagram of a reticulated system, the presence of interaction bonds is schematically indicated. No attempt is made to qualify or describe the form of functional connectedness; there is only the bare statement that such exists.

Energy bonding is a particular type of functional connectedness. The presence of an energy bond will be indicated schematically by a heavy bar between the bonded systems.

$$S_1 \text{——} S_2$$

More than one energy bond can link two given systems. Thus, in general we show one bar for each form of bonding which we wish to consider.

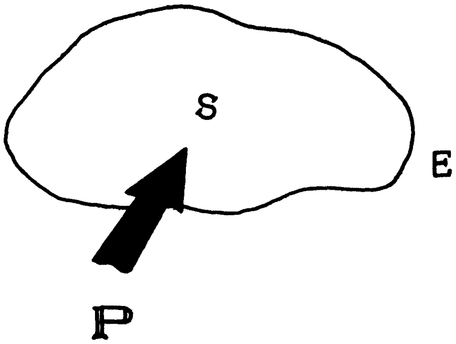


The ever-present possibility of a multiplicity of bonding is a fundamental property of energy which we may call coextensivity or interpenetrability--many forms of energy may occupy the same region of space.

We have established that the analysis of a complex system in its (complex) environment is reliably based upon a general treatment of simple system-to-system interactions. In particular, we shall now focus our attention on a class of material systems called energetic systems and attempt to describe the energy transactions which can occur across the boundaries of such systems.

1. Noncausal Description

Energy transactions across the E-S interface may be defined by a pair of variables which together are a measure of the power or flow rate of energy.



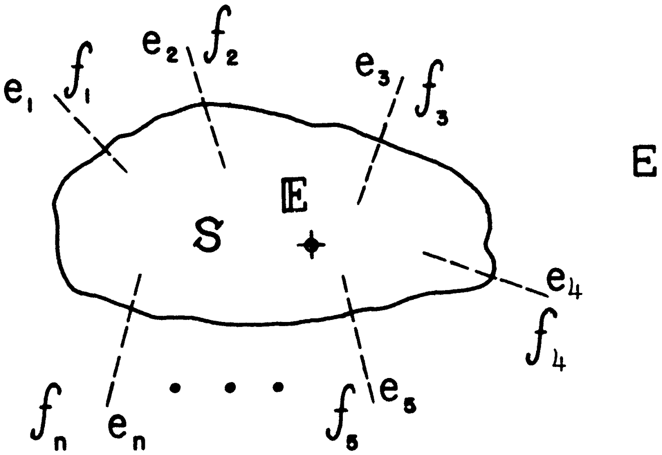
One of the variables is an extensive factor in that its magnitude is dependent upon the extent of the portion of the system entering into the transaction. The other variable is an intensive factor, being a function only of the field in which the system resides. If the two variables are properly chosen, their product will yield the instantaneous power exchanged. The factoring of power into two components is fundamental in mechanics (power = force · velocity) and thermodynamics (power = pressure · rate of volume change). LOTKA relates that attempts have been made in the analysis of social behavior to employ intensities and extensities which combine to yield a quantity analogous to power, although he admits the analogy is a rather loose one.

For our purposes it is convenient to think of the intensive variable as an effort (e) and the extensive variable as a flow (f) so that

$$\text{Power} = \text{Effort} \cdot \text{Flow}$$

$$\mathbb{P} = e \cdot f$$

We then view the system, with which we associate the over-all energy state $\mathbb{E}$, as entering into energy transactions with its environment at a number of localized regions on its boundary surface.



Thus, in the case of the noncausal energy bond reticulation, the quantities e_i and f_i (i = 1, 2, 3, ..., n) are the external variables of the system.

A reticulated energy bond diagram facilitates an over-all grasp of the behavior of a system and an appreciation of the functional connectedness of its elements. More specifically, it imparts an understanding of the transformation and flow of energy within the system and assists in the isolation of the essential energy interactions with the environment.

The manifestation of elusive environmental energy interactions in the behavior of a system has traditionally been the stimulus for innovating discoveries. We are told that NEWTON "discovered" gravity as a result of an apple falling on his head! HUNT gives an account of the discovery of motional impedance by KENNELLY and PIERCE in 1912 which dramatically underscores the importance of environmental interaction in the study of acoustical phenomena.

2. Causal Description

If we wish to attribute a direction of causality to the interaction between two systems, it is imperative that we provide for *bilateral communication*. Employing for the moment the terminology of the previous section, suppose that a flow of energy occurs between two systems, S_1 and S_2. Now, if we wish to endow this interaction with causality, we might be prone to say, for example, that the flow occurs *from* S_1 *to* S_2 by virtue of an effort supplied by S_1. However, this statement is only half true since the effort which S_1 must supply to achieve a given flow is inevitably a function of the *impedance* of S_2 to that flow; hence, a means of "communicating" the nature of this impedance back to S_1 must be provided.

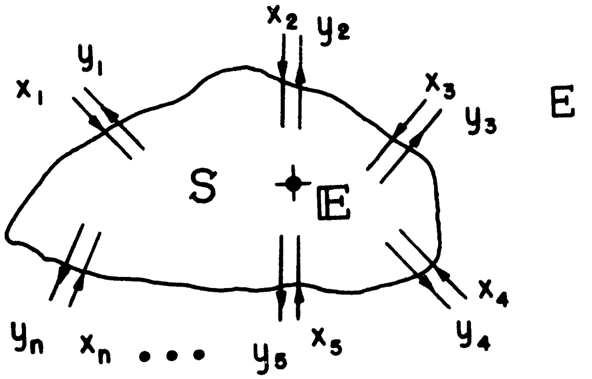
A noncausal *energy bond* can be converted into a causal bond in the form of a *bilateral signal flow*. That is

$$S_1 \overset{e}{\underset{f}{\text{———}}} S_2$$

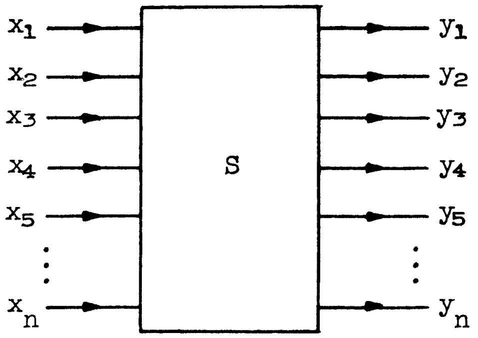
is equivalent to

$$S_1 \underset{y}{\overset{x}{\leftrightarrows}} S_2$$

where the signals x and y are so chosen that $x \cdot y = \mathbb{P}$. Hence, for the general system with n energy transactions occurring across its bounding surface,



We may now imagine a segregation of the n-inputs x_i from the n-outputs y_i and a conceptual deformation of the system such that all the inputs enter at one face while all the outputs leave from the opposite face.



It is convenient to combine the x_i to form an input vector $\mathbb{X} = [x_1, x_2, \dots x_n]$, and the y_i to form an output vector $\mathbb{Y} = [y_1, y_2, \dots x_n]$. Now, for each input-output pair (x_i, y_i) there corresponds a power component $\mathbb{P}_i = x_i y_i$, so that the total power exchanged is

$$\mathbb{P} = \mathbb{X} \circ \mathbb{Y} \qquad = \sum_{i=1}^{n} x_i y_i$$

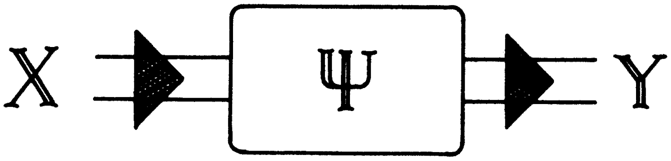
It is interesting to note, in this connection, that power is an invariant under a coordinate transformation. In particular, let the matrix of the transformation be T. Then, LE CORBEILLER demonstrates that either $\mathbb{X}$ or $\mathbb{Y}$ will be contravariant, i.e., it will transform like the coordinates according to the matrix T; while the other will be covariant, i.e., it will transform according to the transpose of T^{-1}. Suppose it is $\mathbb{X}$ that is covariant. Then the power is given by the matrix multiplication

$$\mathbb{P} = \mathbb{X} \cdot \mathbb{Y}_t$$

where $\mathbb{Y}_t$ denotes the transpose of $\mathbb{Y}$. In the new coordinate frame

$$\mathbb{P} = [(T^{-1})_t \mathbb{X}]\, T\mathbb{Y}_t = \mathbb{X}\, T^{-1}\, T\mathbb{Y}_t = \mathbb{X}\ \mathbb{Y}_t = \mathbb{P}$$

Finally, the system S is conceptually replaced by a module which performs the operation on $\mathbb{X}$ which is required to yield up $\mathbb{Y}$.



The functional dependency $\mathbf{Y} = \mathbf{\Psi}(\mathbf{X})$ is of a most general form such that the entire history of $\mathbf{X}$ is scanned to yield a present value of $\mathbf{Y}$. Therefore, it is applicable to the analysis of all processes in which the system S might be involved, and in particular to transition processes from one steady-state condition to another.

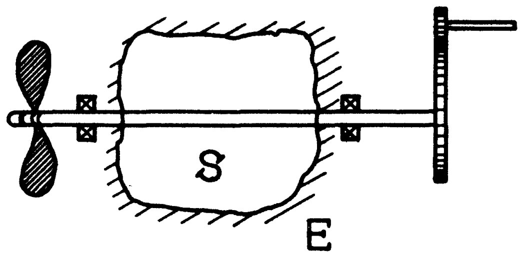
For a purely steady-state analysis, a simpler static function $\mathbf{\Phi}$ is applicable, yielding the present value of $\mathbf{Y}$ corresponding to a present value of $\mathbf{X}$.

Thus, we see that in the case of a bilateral signal flow reticulation the variables of the system are the vectors $\mathbf{X}$ and $\mathbf{Y}$, or, more precisely, their respective components.

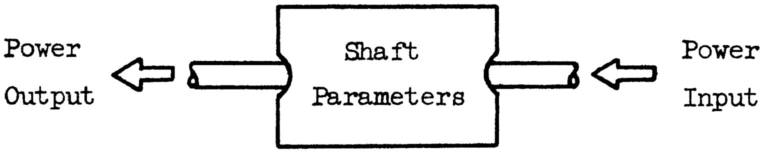
Contained in the function $\mathbf{\Psi}$ are the intrinsic properties of the system in the form of a set of parameters. In the most general case the parameters may vary with time and with the environment of the system. Thus we may display the scheme:

System Characteristics
- Variables
  - $\mathbf{X}$ & $\mathbf{Y}$
    Input (exogenous)
    Output (endogenous)
  - $\mathbf{e}$ & $\mathbf{f}$
    Effort
    Flow
- Parameters
  - Geometrical
    Diameter,
    Length, etc.
  - Material
    Elastic and inertial
    Properties, etc.

As an example, consider the propeller shaft in the turbo-drive propulsion system discussed in Part I.

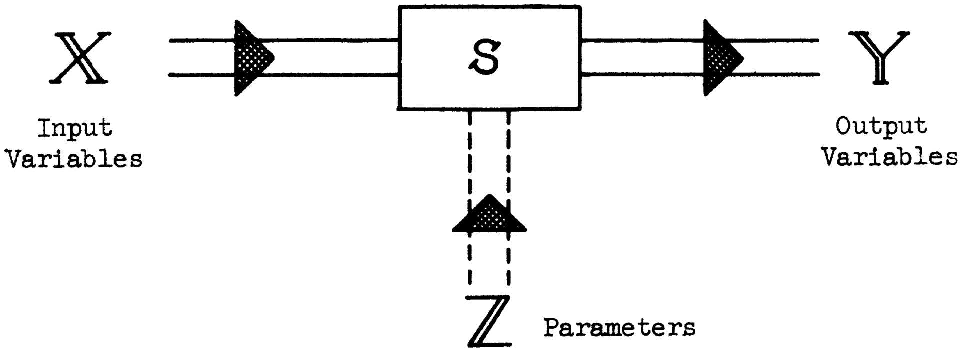


First the system is suitably isolated from its environment; it is then conceptually transfigured into a module possessing the properties of the shaft.



A practical analysis of a system, such as the propeller shaft in the above example, is directed either toward a determination of behavior for a given design, or the design which will achieve a given behavior, in both cases the environment being taken as fixed. Thus, it is sometimes convenient to view the parameters of the system also as "inputs," redrawing the module as shown below:

System S



Hence, problems of the first type are involved with finding $\mathbb{Y}$ as a function of $\mathbb{X}$ for a given $\mathbb{Z}$, while those of the second type involve finding $\mathbb{Z}$ as a function of $\mathbb{Y}$ for a given $\mathbb{X}$.

Background Reading

(1) BRØNSTED, J. N. *Principles and Problems in Energetics,* Chapters I and II.

Chapter I imparts some feeling for the relationship between Energetics and the classical theories of thermodynamics. Chapter II demonstrates the inadequacy of the classical work concept and delineates more carefully the requirements of a generalized *work principle*.

(2) LOTKA, A. J. *Elements of Mathematical Biology,* pp.280-286; 303-305.

One is led to the belief that there are characteristics of social systems which are, at least in an *abstract* sense, describable *energetically*.

(3) DE GROOT, S. R. *Thermodynamics of Irreversible Processes,* pp.1-9.

The author points out, in connection with the effort-flow concept, that there are many examples of "cross-effects"-- coupling between two processes, such as in the thermo-electric effect, wherein an effort of one form produces, in addition to its corresponding flow, a flow of another form. Thus, one's attention is directed towards the problem of describing such a coupling by way of energy bonds. A presentation of ONSAGER's theory provides some understanding as to how the problem might be handled.

(4) HUNT, Frederick V. *Electro-Acoustics,* p.96.

See particularly the interesting account of the discovery of *motional impedance* by KENNELLY and PIERCE.

(5) TRIMMER, J. D. *Response of Physical Systems,* pp.98-103.

The ideas of *conjugate variables* and *impedances* are briefly introduced.

(6) ZWIKKER, C. *Physical Properties of Solid Materials,* pp.72-73.

An appreciation is imparted of the parametric description of a system in the form of a set of *compliances* or *rigidities,* for the static case, and *conductances* or *resistances* in the stationary case.

(7) LE CORBEILLER, P. Matrix Analysis of Electrical Networks, pp.59-61.

The invariance of power under a coordinate transformation is demonstrated for a system residing in a potential force field. The distinction between contravariant quantities and covariant quantities is made.

IV. The Continuity of Energy

A. Introduction and Historical Background

Isaac NEWTON's view of the physical universe was that of a system of mass points whose interactions and over-all behavior were describable in terms of certain basic laws of mechanics, these laws being interpreted for a specific particle as the total differential equations of motion. NEWTON and his followers attempted to extend this explanation of reality to encompass all forms of interaction. However, their mechanistic explanation of light, being, as it were, founded on particle interaction, left much to be desired.

Man's appreciation of the universe was dramatically broadened with the advent of James Clerk MAXWELL's use of field concepts for light and electricity. Here was a theoretical structure of sufficient scope to serve as a skeleton for the analysis of the electromagnetic radiation phenomena which had eluded physicists of the Newtonian school. Although he did not view fields specifically as "energy fields," MAXWELL was certainly aware of the energetic aspect of fields of all types.

Albert EINSTEIN sums up the effect which MAXWELL had on physics in the following quotation:

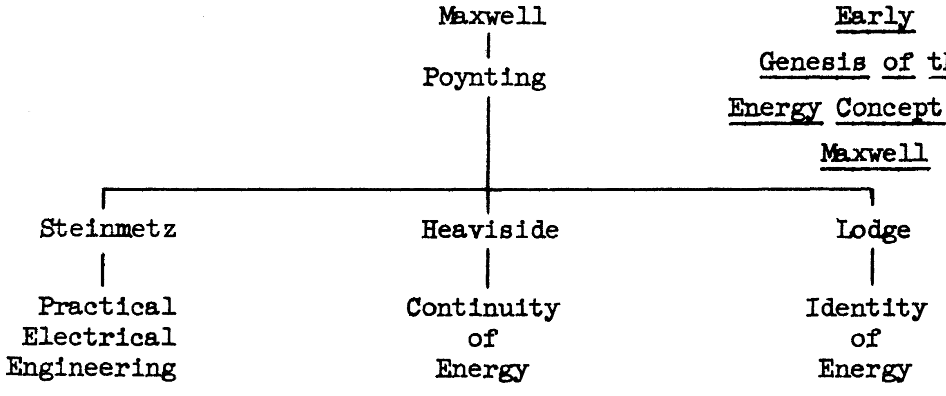
> Before Clerk Maxwell people conceived of physical reality--in so far as it is supposed to represent events in nature--as material points, whose changes consist exclusively of motions which are subject to total differential equations. After Maxwell they conceived physical reality as represented by continuous fields, not mechanically explicable, which are subject to partial differential equations. This change in the conception of reality is the most profound and fruitful one that has come to physics since Newton; but it has at the same time to be admitted that the program has by no means been completely carried out yet. The successful systems of physics which have been evolved since rather represent compromises between these two schemes, which for that very reason bear a provisional, logically incomplete character, although they may have achieved great advances in certain particulars.

John Henry POYNTING extended MAXWELL's formulation to the consideration of energy transport in an electrical network. The scope of POYNTING's analysis is best described in his own words:

> A space containing electric currents may be regarded as a field where energy is transformed at certain points into

> the electric and magnetic kinds by means of batteries, dynamos, thermoelectric actions, and so on, while in other parts of the field this energy is again transformed into heat, work done by electromagnetic forces, or any form of energy yielded by currents. Formerly a current was regarded as something travelling along a conductor, attention being chiefly directed to the conductor, and the energy which appeared at any part of the circuit, if considered at all, was supposed to be conveyed thither through the conductor by the current. But the existence of induced currents and of electromagnetic actions at a distance from a primary circuit from which they draw their energy, has led us, under the guidance of FARADAY and MAXWELL, to look upon the medium surrounding the conductor as playing a very important part in the development of the phenomena. If we believe in the continuity of the motion of energy, that is, if we believe that when it disappears at one point and reappears at another it must have passed through the intervening space, we are forced to conclude that the surrounding medium contains at least a part of the energy, and that it is capable of transferring it from point to point.

On the basis of POYNTING's formulation, Charles Proteus STEINMETZ developed a regimented theory for the practical analysis and design of circuits and laid the foundations for electrical engineering. Oliver HEAVISIDE continued the thread of POYNTING's work to a generalized statement of energy continuity, while Sir Oliver LODGE discussed the identity and conservation of both matter and energy. Thus, the concept of energy-matter conservation in space as well as time became firmly implanted in the structure of the theory of Energetics.



Early Genesis of the Energy Concept from Maxwell

In parallel with the line of development sketched above was that of the classical theory of thermodynamics, which generated an energy concept of a rather divergent nature. From the fundamental experiments of Julius Robert MAYER, James Prescott JOULE, and William Thomson (Lord KELVIN) there emerged the principles of energy conservation, energy transformation, and the equivalence of heat and work--ideas upon which LODGE drew heavily. However, as has been previously stated, the axiomatic structure of thermodynamics is applicable only to systems whose boundaries are of limited permeability. Rudolf CLAUSIUS made sweeping statements concerning the energy and entropy of the entire universe, but only on the basis of the assumption that the universe could be considered an isolated system! Regarding the conservation of energy, HEAVISIDE had this to say:

> The principle of the continuity of energy is a special form of that of its conservation. In the ordinary understanding of the conservation principle it is the integral amount of energy that is conserved, and nothing is said about its distribution or its motion. This involves continuity of existence in time, but not necessarily in space also.
>
> But if we can localise energy definitely in space, then we are bound to ask how energy gets from place to place. If it possessed continuity in time only, it might go out of existence at one place and come into existence simultaneously at another. This is sufficient for its conservation. This view, however, does not recommend itself. The alternative is to assert continuity of existence in space also, and to enunciate the principle thus:--
>
> When energy goes from place to place it traverses the intermediate space.
>
> This is so intelligible and practical a form of the principle, that we should do our utmost to carry it out.

But one now has the right to inquire, as did MAXWELL, "If something is transmitted from one particle to another at a distance, what is its condition after it has left the one particle and before it has reached the other?" It was indeed necessary to endow the void with certain material properties in order to conceive of a transfer of energy through it. For this purpose the ether was contrived--a substance to permeate all space through which energy might pass. That MAXWELL was unable to conceive of an energy transfer apart from an intervening medium is clearly indicated by the following quotation:

> In fact, whenever energy is transmitted from one body to another in time there must be a medium or substance in which

> the energy exists after it leaves one body and before it reaches the other, for energy, as Torricelli remarked, "is a quintessance of so subtle a nature that it cannot be contained in any vessel except the inmost substance of material things."

An ether was also essential in the eyes of POYNTING, LODGE, and HEAVISIDE, for these men felt that the transformation of energy was always accompanied by a transfer, and vice versa. This was the very essence of the principles of continuity and identity which they propounded.

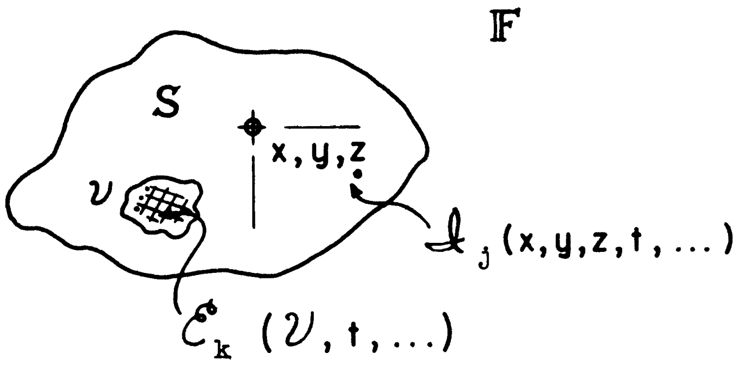
As man's understanding of the physical universe was extended his conception of the ethereal substance was correspondingly altered. A modern viewpoint is submitted by EINSTEIN by way of conclusion to his essay, "Relativity and the Ether:"

> We may sum up as follows: According to the general theory of relativity space is endowed with physical qualities; in this sense, therefore, an ether exists. In accordance with the general theory of relativity space without an ether is inconceivable. For in such a space there would not only be no propagation of light, but no possibility of the existence of scales and clocks, and therefore no spatio-temporal distances in the physical sense. But this ether must not be thought of as endowed with the properties characteristic of ponderable media, as composed of particles the motion of which can be followed; nor may the concept of motion be applied to it.

For our present purposes the universe will be viewed as a field--a heterogenous, anisotropic continuum, i.e., its properties are variable over space and polarized with respect to orientation. The ether is then a conceptual artifice which we shall dispense with; indeed, we shall take the even stronger position that energy may exist in and of itself, requiring no material vessel in which to be stored or transported.

B. Properties of the Field

If we think of all the universe as constituted from a relatively small number of basic particles (about thirty), then the only difference between the air, a wall, a floor, a desk, etc., is the density and distribution of these particles. Then consider a system S residing in a field F:



We describe the system in terms of variables and parameters which fall into two categories: (i) Intensive quantities $\mathcal{I}_j(x,y,z,t,\ldots)$ which are characteristics of the field F; (ii) Extensive quantities $\mathcal{E}_k(\mathcal{V},t,\ldots)$ which depend also on the extent ($\mathcal{V}$) of the system involved. However, the existence of extensive quantities cannot be reconciled directly with the field description which we seek; it is thus convenient to define specific extensities or field densities which are the densities of the extensive quantities $\mathcal{E}_k$. We define

$$\epsilon_k = \lim_{\mathcal{V}\to 0} \mathcal{E}_k/\mathcal{V} = \epsilon_k(x,y,z,t,\ldots)$$

Thus, the ϵ_k are point quantities and are conformable to the field description. Perhaps the most familiar example of a specific extensity is the mass density, ρ, where with mass, M, in a volume, $\mathcal{V}$:

$$\rho = \lim_{\mathcal{V}\to 0} M/\mathcal{V}$$

C. Generalized Continuity Equation

The statement of the principle of mass conservation as a field equation is familiar to us,

$$\operatorname{div}(\rho \vec{V}) + (\partial\rho/\partial t) = 0$$

We know that $\operatorname{div}(\rho \vec{V})$ measures the convective transfer of mass densities away from a point in the field while $(\partial\rho/\partial t)$ is the time rate of change in the local density. Setting the right-hand side equal to zero assumes there are neither sources or sinks of matter in the field, tantamount to common parlance "mass is neither created nor destroyed." More generally, however, in the light of modern physics and mass-energy equivalence, we might insert a source term σ and write

$$\mathrm{div}(\rho \vec{V}) + (\partial\rho/\partial t) = \sigma$$

Now, it is quite appropriate that we similarly demand the continuity of any and all of the specific extensities ϵ_k. Therefore, we may state the generalized equation of continuity as follows:

$$\mathrm{div}(\epsilon \vec{V}) + (\partial\epsilon/\partial t) = \sigma$$

where $\epsilon \vec{V}$ is the convective transport of ϵ and $\vec{V}$ is the appropriate transport velocity. Sometimes, in order to avoid the necessity of measuring V it is convenient to define a ϵ-flux vector, $\vec{\mathcal{F}}$, as

$$\vec{\mathcal{F}} = \epsilon \vec{V}$$

so that we can write:

$$\boxed{\mathrm{div}(\vec{\mathcal{F}}) + (\partial\epsilon/\partial t) = \sigma}$$

D. Continuity of Energy and the Generalized Poynting Vector.

MAXWELL dealt specifically with the localized energies of a field. For example, he attributed to an electrostatic field a distributed potential energy proportional to the product of the potential ψ and the electrical displacement e. He likewise defined magnetic and electrokinetic energies in terms of distributed quantities--energy densities, as it were.

The discussion of energy on a localized basis allows one to be most specific, and, in general, uncompromising in the analysis of a particular system. Classical modes of analysis always require, as has been repeatedly pointed out, a "cloaking" of the system with a quasi-impenetrable veil, and generally demand that energy be integrated over the domain of the system before any acceptable statements concerning its conservation or transformation may be made. Thus, we might say that the conception of energy as a localized or distributed quantity permits the fullest possible exploitation of the available information concerning the nature of the system, its bounding surface, and its conceivable reticulations. Hence, following HEAVISIDE's logic directly, we shall state the equation of energy continuity for a field:

$$\boxed{-\operatorname{div}\vec{\mathbb{p}} = (\partial\epsilon/\partial t) + \rho_d}$$

where $\vec{\mathbb{p}}$ is the generalized Poynting Vector, ϵ is the local energy density, and ρ_d is the rate of energy loss or dissipation. This statement will be the foundation for our analyses of energetic systems.

E. Continuity of Entropy

It is appropriate at this point to introduce parenthetically and without development a conjugate continuity equation for entropy.

$$-\operatorname{div}\vec{\mathbb{S}} = (\partial\phi/\partial t) + \mathfrak{s}$$

where $\vec{\mathbb{S}}$ = entropy flux, ϕ = entropy density, and $\mathfrak{s}$ = entropy sink. This equation is most important in that it permits the analysis of a system which is exchanging entropy and information with its environment rather than, or perhaps in addition to, energy. Biological systems are notable examples of systems to which it is applicable. We shall discuss the relation of this equation to energy continuity somewhat later on.

Background Reading

(1) EINSTEIN, A. Essays in Science.

"Clerk Maxwell's Influence on the Evolution of the Idea of the Physical Universe"

"The Problem of Space, Ether, and the Field of Physics"

"Relativity and the Ether"

This group of essays imparts a very broad appreciation of the development of the physical concepts relevant to Part IV from Newtonian mechanics, through MAXWELL's field theory, to the modern theories of relativity and quantum mechanics.

(2) POYNTING, J. H. On the Transfer of Energy in the Electromagnetic Field, Philosophical Transactions of the Royal Society of London for the Year 1884, Vol. 175, pp.343-361.

The manner in which energy is transported from point to point in a space carrying electrical currents and is transformed at various points by way of batteries, motors, etc., is analyzed.

(3) HEAVISIDE. O. Electro-magnetic Theory, pp.73-77.

The principle of energy continuity is stated by way of an extension to POYNTING's conception of energy transport.

(4) LODGE, O. On the Identity of Energy, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, Vol. XIX, January-June (1885) pp.482-487.

As another extension to POYNTING's energy transport concept, the idea of the identification of energy, as it is transferred from one point to another, is discussed.

(5) MAXWELL, J. C. Electricity and Magnetism, Vol. III, pp.270-274; 493.

In the first reference MAXWELL develops expressions for the various energies of a field. In the second he makes a concluding statement concerning what he feels to be the nature of the ether.

(6) SOMMERFELD, A. Thermodynamics and Statistical Mechanics, pp.152-153.

An equation for entropy continuity is derived from that for energy continuity for the special case of a homogeneous, isotropic solid.

(7) PRIGOGINE, I. Thermodynamics of Irreversible Processes, pp.32-34.

A general equation of entropy continuity is developed similar to the one stated above.

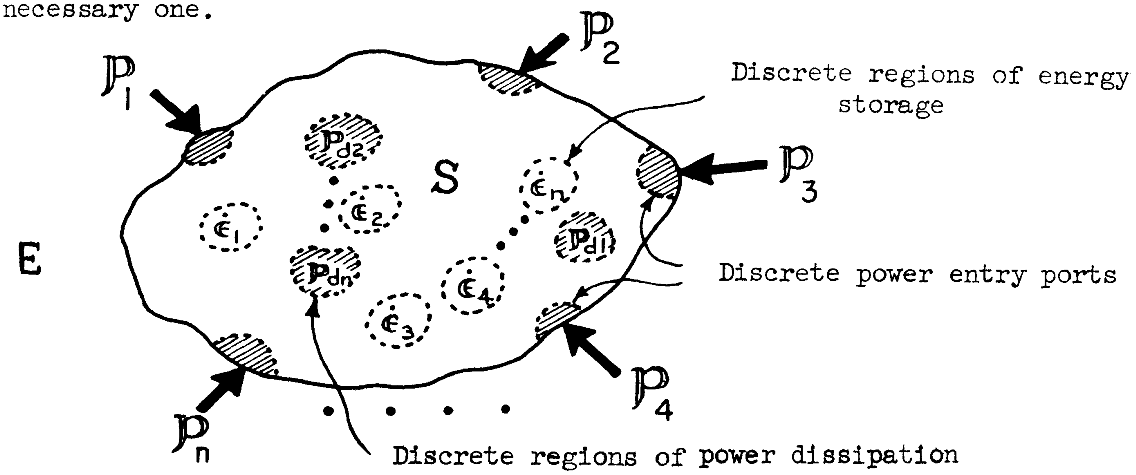
V. Energy Ports and Power Bonds

A. Introduction

We have come to view the universe as a field in which resides the system S. We are now in a position to require that energy transactions between S and E and within S be subject to the general equation of energy continuity

$$- \operatorname{div} \vec{\mathbb{P}} \equiv \frac{\partial \epsilon}{\partial t} + \mathcal{P}_d$$

The purpose of Part V is to reticulate the continuity equation for application to the analysis of energetic systems. This amounts to assuming: (1) that the S-E interface is permeable to the passage of energy only over a relatively few areas of restricted extent--energy ports as we shall call them; (2) that the energy storage function of the system is not distributed continuously throughout its volume, but is rather lumped in discrete localities or regions; (3) that the dissipative property of the system is also confined to discrete regions. Although this is admittedly an idealization, it is a very practical and necessary one.



Reticulation of Energy Transactions,
Storages, and Dissipations

B. Application of the Divergence Theorem

The Divergence Theorem will now be applied to the continuity equation as stated above in order that it be converted to a form amenable to reticulation. The Theorem states

$$\int_{\mathcal{A}} \vec{\mathbb{F}} \cdot \vec{n} d\mathcal{A} = \int_{\mathcal{V}} \text{div } \vec{\mathbb{F}} d\mathcal{V}$$

where $\mathcal{A}$ is the closed surface bounding the volume $\mathcal{V}$ and n is the unit outward normal. This single theorem is perhaps the most profound in all applied mathematics since from it every other theorem may be derived.

Substituting the generalized Poynting vector, $\vec{\mathbb{P}}$, for the generic vector, $\vec{\mathbb{F}}$, and letting $\mathcal{V}$ be the volume of S and $\mathcal{A}$ the area of its bounding surface

$$\int_{\mathcal{A}} \vec{\mathbb{P}} \cdot \vec{n} d\mathcal{A} = \int \text{div } \vec{\mathbb{P}} d\mathcal{V}$$

Performing the volumetric integration on all terms in the continuity equation, and substituting from the above we thus obtain

$$-\int_{\mathcal{A}} \vec{\mathbb{P}} \cdot \vec{n} d\mathcal{A} = \int_{\mathcal{V}} \left(\frac{\partial \mathcal{E}}{\partial t} + \mathcal{P}_d\right) d\mathcal{V}$$

$$= \int_{\mathcal{V}} \frac{\partial \mathcal{E}}{\partial t} d\mathcal{V} + \int_{\mathcal{V}} \mathcal{P}_d d\mathcal{V}$$

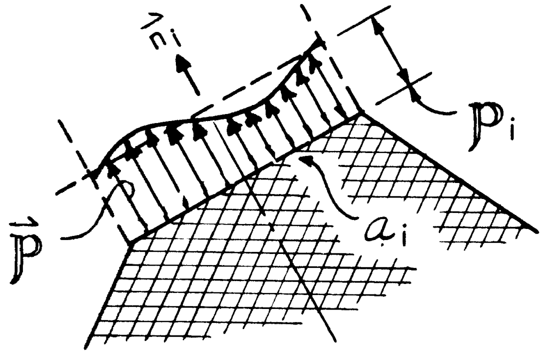
The continuity equation now stands ready to be reticulated in the manner we desire.

C. Reticulation of the Energy Influx

We consider first the convective transfer of energy across the system boundary; the assumption is made that there are l areas $\mathcal{A}_i$ over which the S-E interface is permeable to energy flow. It is further assumed that each of the $\mathcal{A}_i$ is of sufficiently limited extent that $\vec{\mathbb{P}}$ may be considered everywhere parallel to the associated unit normal $\vec{n}_i$, and that nonuniformities in the magnitude of $\vec{\mathbb{P}}$ may be absorbed by an appropriate averaging operation such that

$$\int \vec{\mathbb{P}} \cdot \vec{n}_i d\mathcal{A}_i = \mathbb{P}_i \mathcal{A}_i$$

Thus, the boundary of S is no longer viewed as a shapeless bag, but rather a multifaceted surface, each facet corresponding to an energy port.



Now, for each port the product $\mathbf{p}_i\, a_i$ is the power leaving through that port, $\mathbb{P}_i$. Hence, the total power flux through the boundary of the system is

$$\sum_{i=1}^{l} \mathbf{p}_i\, a_i = \sum_{i=1}^{l} \mathbb{P}_i$$

D. Reticulation of Energy Storage

It would appear sensible to assume that energy storage is a function of a relatively few (m) localized regions in S rather than being uniformly distributed throughout its volume. Hence, when the integration

$$\int_{\mathcal{V}} \frac{\partial \mathcal{E}}{\partial t}\, d\mathcal{V}$$

is performed over the entire volume of S there will only be a discrete number of regions $\mathcal{V}_j$ in which

$$\frac{\partial \mathcal{E}}{\partial t} \neq 0$$

Thus

$$\int_{\mathcal{V}} \frac{\partial \mathcal{E}}{\partial t}\, d\mathcal{V} = \sum_{j=1}^{m} \int_{\mathcal{V}_j} \frac{\partial \mathcal{E}}{\partial t}\, d\mathcal{V}_j$$

Now, a reticulation of the energy storage in S will have meaning only if the volumes $\mathcal{V}_j$ in which storage is supposed to occur are fixed in size and disposition. Assuming this to be the case, we may carry the time differentiation outside the integral and change it from partial to total giving

$$\int_{\mathcal{V}} \frac{\partial \mathcal{E}}{\partial t} d\mathcal{V} = \sum_{j=1}^{m} \frac{d}{dt} \int_{\mathcal{V}_j} \mathcal{E} d \mathcal{V}_j$$

But, we now have the right to define

$$\mathbb{E}_j = \int_{\mathcal{V}_j} \mathcal{E} d \mathcal{V}_j$$

as the instantaneous energy stored in each little storage region $\mathcal{V}_j$. Hence, we finally obtain

$$\int_{\mathcal{V}} \frac{\partial \mathcal{E}}{\partial t} d \mathcal{V} = \sum_{j=1}^{m} \frac{d\mathbb{E}_j}{dt}$$

E. Reticulation of Energy Dissipation

Following the same logic, it is a plausible assumption that energy dissipation will occur only in a discrete number (n) of restricted regions $\mathcal{V}_k$, i.e., the resistive or dissipative property of S is lumped in the regions $\mathcal{V}_k$ rather than being distributed throughout $\mathcal{V}$.

It is therefore evident that

$$\int_{\mathcal{V}} p_d d \mathcal{V} = \sum_{k=1}^{n} \int_{\mathcal{V}_k} p_d d \mathcal{V}_k$$

with the definition

$$\int_{\mathcal{V}_k} p_d d \mathcal{V}_k = (\wp_d)_k$$

we finally obtain

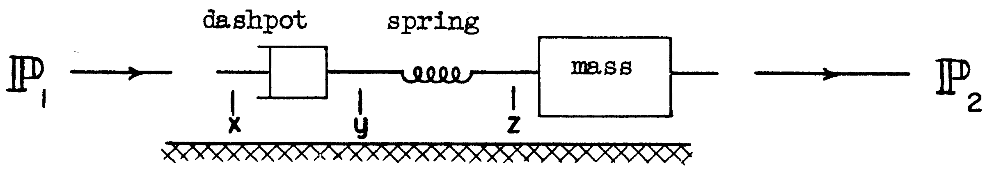
$$\int_{\mathcal{V}} p_d d \mathcal{V} = \sum_{k=1}^{n} (\wp_d)_k$$

F. The Reticulated Equation of Energy Continuity

Combining the above results we may write the *reticulated equation of energy continuity as follows*:

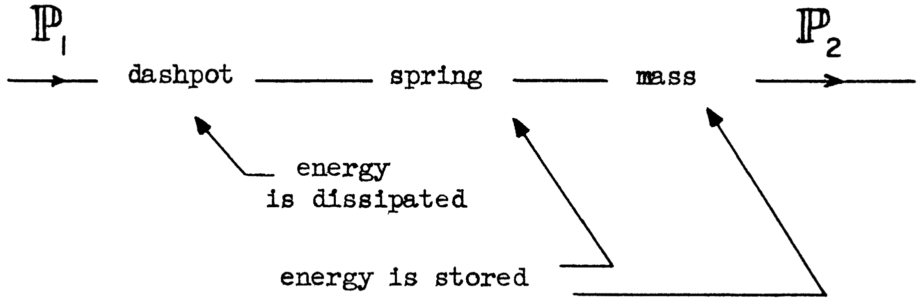
$$-\sum_{i=1}^{l} \mathbb{P}_i = \sum_{j=1}^{m} \frac{d\mathbb{E}_j}{dt} + \sum_{k=1}^{n} (\wp_d)_k$$

This equation makes the irrefutable statement that a net flux of energy into a system is either stored or dissipated, and on it we may found the practical analysis of any energetic system. However, the identification and calculation of the $\mathbb{P}_i$, the $\mathbb{E}_j$, and the $(\wp_d)_k$ in terms of the variables and parameters of an actual system is no trivial task. A simple example will serve to reveal the first major pitfall which must be avoided.



$\xi = y - x$

$\eta = z - y$



We intend to analyze a spring-mass-dashpot system in which the spring and dashpot are nonlinear, and we wish to account for the relativistic variation in the momentum of the mass. In particular, suppose we are given the following characteristic relations:

$$\text{Spring force} = e_s = e_s(\eta)$$

$$\text{Damping force} = e_d = e_d(\dot{\xi})$$

$$\text{Momentum of mass} = p = p(\dot{z})$$

Since energy is stored in the mass and the spring, and is dissipated in the dashpot, we may say with assurance

$$\mathbb{P}_1 - \mathbb{P}_2 = \frac{d}{dt}[\mathbb{E}_s + \mathbb{E}_m] + \mathcal{P}_d$$

$$= \frac{d}{dt}\left[\int e_s d\eta + \int \dot{z} dp\right] + e_d\dot{\xi}$$

In this case $\mathcal{P}_d$ is self-evident; however, one might be prone to write instead of $\mathbb{E}_s$ and $\mathbb{E}_m$ what we might label the complementary energies,

$$\mathbb{E}_s^* = \int \eta\, de_s \quad ; \quad \mathbb{E}_m^* = \int p\, d\dot{z}$$

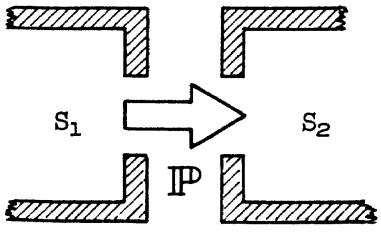
It is noteworthy that $\mathbb{E}_s = \mathbb{E}_s^*$ and $\mathbb{E}_m = \mathbb{E}_m^*$ only for a linear spring and constant mass. Thus, it is imperative that extreme caution be exercised in evaluating the "energies" of a system so that the inclusion of "incorrect" energy terms may be avoided. Needless to say, the continuity equation is valid only if the energy terms are properly evaluated.

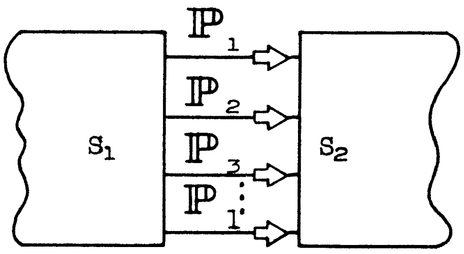
A similar difficulty can arise in connection with the power flows $\mathbb{P}_i$. Power is carried across the system boundary by transmission links--shafts, ducts, electrical conductors, waveguides, etc. We have previously stated that the power state of a transmission link may be indicated by the product of an intensive variable (effort--e) and an extensive variable (flow--f), which tacitly assumes that two such variables may be identified for a given transmission link. For the description of the power flow through a shaft, for example, we would be prone to pick the torque as the effort and the angular velocity as the flow. But what is the torque of a shaft? Indeed, a shaft possesses neither a single, characteristic quantity "the torque," nor a single angular velocity common to all its parts! Granted that we could integrate the moment of the shear stress (τ) over the cross-section of the shaft and get a torque $M = \int_A \tau \, r dA(r)$, and we could also calculate a mean angular velocity $\bar{\omega} = \dfrac{\int_A \omega(r) dA(r)}{\int_A dA(r)}$, but who is to guarantee that the product $M \cdot \bar{\omega}$ will yield the true power transmitted? Our only recourse is to calculate (or measure) one of the variables, say the torque, and then to assume an angular velocity $\tilde{\bar{\omega}}$ ($\not\equiv \bar{\omega}$) such that $M \cdot \tilde{\bar{\omega}}$ does in fact give the power. A course of action similar to this must be taken in the case of any real (i.e., nonideal) transmission link.

We shall suppose that it will always be possible to find a pair of effort and flow variables such that, for a given system,

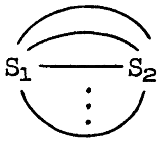
$$\mathbb{P}_i = e_i \cdot f_i \quad ; \quad i = 1, 2, 3, \cdots, l$$

The energy coupling between a system and its environment is often rather elusive. In particular, a single transmission link may appear to be the medium of exchange for several, or perhaps even an infinite number of energies. For example, any small region of space is energetically coupled to every other region by way of a spectrum of electromagnetic radiations. It behooves us, therefore, to allot one energy port to each transaction; hereafter, we shall speak of an energetic interaction between two systems as a power bond. Thus, the power bonding between a system and its environment, or between two systems S_1 and S_2 is reticulated as sketched below into l separate bonds.





Or, more simply,



where one bond is allotted to each energetic interaction.

Once the power bonding between S_1 and S_2 has been fully reticulated, and each bond has been described in terms of an effort X flow product, we are in a position to define the vectors:

$$\mathbb{e} = [e_1, e_2, e_3, \ldots, e_l]$$
$$\mathbb{f} = [f_1, f_2, f_3, \ldots, f_l]$$

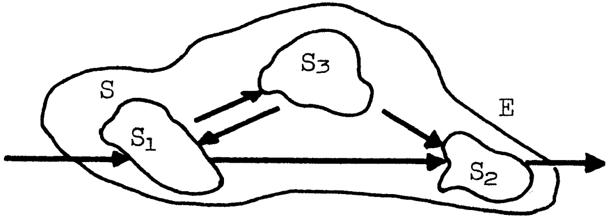
Thus, the total power transacted is given by the matrix product

$$\mathbb{P} = \mathbb{e}_t \cdot \mathbb{f} = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_l \end{bmatrix} [f_1, f_2, f_3, \ldots, f_l]$$

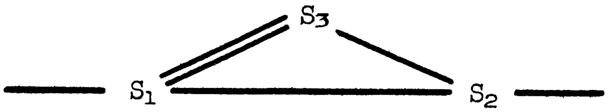
$$= \sum_{i=1}^{l} e_i \cdot f_i$$

G. Power Bonds

The schematic representation of a reticulated energetic system which we have adopted shows the elements as linked by heavy bars, each bar denoting a power bond. Thus, for a two-port system which is reticulated into three multiported elements,



the simple bond diagram would appear as follows



We note immediately the similarity between such a diagram and a <u>chemical</u> <u>bond</u> <u>diagram</u>. Indeed, the similarity is by no means superficial the mechanism by which a chemical bond is created between two atoms is closely analogous to the formation of a power bond linking two systems, the atomic valency in the former case corresponding to the energetic portality in the latter. Thus, we have the chemists to thank for some of the essential ideas incorporated into our schematic representation. In particular, KEKULE and FRANKLAND were early proponents of the bond diagram.

Causality of Power Bonds

Causality implies the existence of two variables, one independent and the other dependent--such as in a mathematical relationship y = f(x) wherein a y-value inevitably follows once a x-value is specified.

$$\left.\begin{matrix}\text{From}\\ \text{Independent}\end{matrix}\right\} \longrightarrow \left\{\begin{matrix}\text{To}\\ \text{Dependent}\end{matrix}\right.$$

$$f(x) \quad = \quad y$$

Suppose, for example, that

$$y = ax + b = f(x)$$

we can also write, in this case, the inverse relation

$$x = (y/a) - (b/a) = f^{-1}(y)$$

Thus, there is no indication of causality inherent in the sign of equality; rather, by convention an equation is generally written so that the dependent variable is on the left, thus implying a right to left causality. No ambiguity results, however, if we were to indicate the functional dependency of y on x by writing

$$x \longrightarrow \boxed{f} \longrightarrow y$$

Then, applying the inverse

$$x \longrightarrow \boxed{f} \longrightarrow y \longrightarrow \boxed{f^{-1}} \longrightarrow x$$

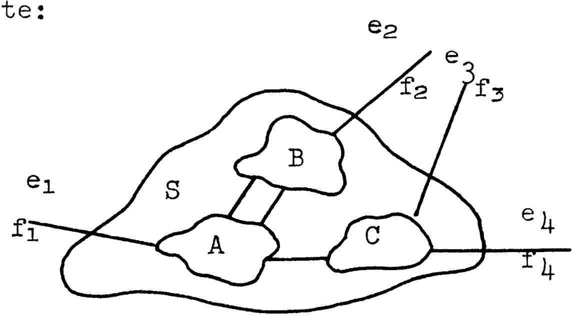
Such a representation emphasizes the signal sense and de-emphasizes the nature of the functional dependence and the form of the "signals" x and y.

We are reminded of the care with which the chemist endows his equations with causality, this being due to the fact that the direction in which a reaction proceeds is so intimately dependent upon the ambient conditions. Thus, for example, the direction of causality in the equation

$$2H_2 + O_2 \rightleftarrows 2H_2O + \text{heat}$$

is partly determined by the ambient temperature. The reverse (right to left) reaction predominates at extremely high temperatures, whereas the forward reaction predominates at lower temperatures.

For our purposes it is imperative that a direction of causality be imparted to an energetic exchange since no quantitative analysis of any form is possible until this is done. Once a power bond has been described in terms of an effort-flow couple then it may be endowed with causality. In the case of a system communicating energetically with its environment either the effort or the flow may be viewed as determined by the environment, i.e., as <u>independent</u>, so that the other variable is looked upon as <u>dependent</u>. Consider, for example, an ideal fluid system in the steady-state:



Suppose that for this system the pressure is determined by the environment and is held constant (but not necessarily uniform) over the entire boundary of the system. Assuming the fluid to be incompressible, it would be natural to set

$$e_i = \text{pressure in lb./ft}^2$$

$$f_i = \text{flow rate in ft}^3\text{/sec}$$

in order to describe the four power bonds between S and E. Since the pressure is environmentally determined

$$e_1 = E_1 = \text{constant}$$

$$e_2 = E_2 = \text{constant}$$

$$e_3 = E_3 = \text{constant}$$

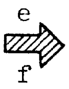
$$e_4 = E_4 = \text{constant}$$

The continuity equation states, for this example

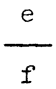
$$\sum \mathbb{P}_i = \sum e_i f_i = \sum E_i f_i = 0$$

since there cannot be internal energy dissipation or storage. In addition, each of the $\mathbb{P}_i$ is constant so that each of the f_i is also constant. Thus, the efforts and flows associated with the internal bonds must be constant.

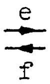
We have stated that an energetic interaction is endowed with causality if it is conceived as a bilateral signal flow. Thus, a power bond



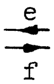
, or simply,



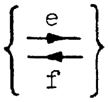
becomes



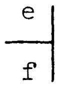
or

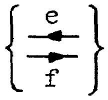


The assignment of causality to a bond is equivalent to adding a single bit of information to the noncausal bond. Hence, it is theoretically possible to accomplish this addition with a single stroke, thus:

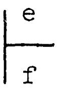


is equivalent to

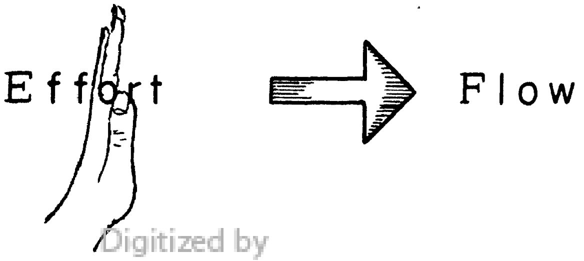




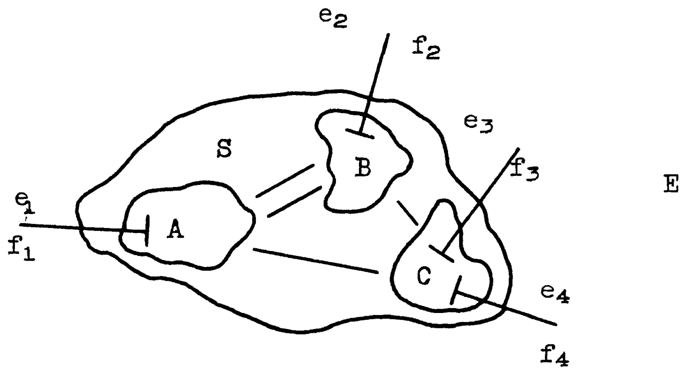
is equivalent to



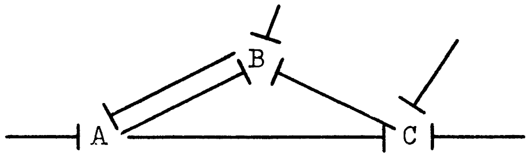
A useful mnemonic is the association of the flow variable, f, with a direction parallel to or along the bond, and the effort variable, e, with the transverse stroke.



Thus, in the four-port fluid system we would indicate the fact that the pressure is given on the boundary as follows:



Assigning causality arbitrarily to the internal bonds, we may represent the system completely in the following succinct form:

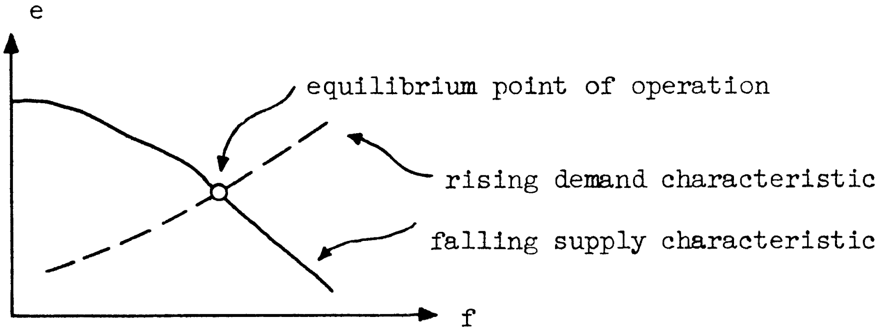


Equilibrium Power Transfer

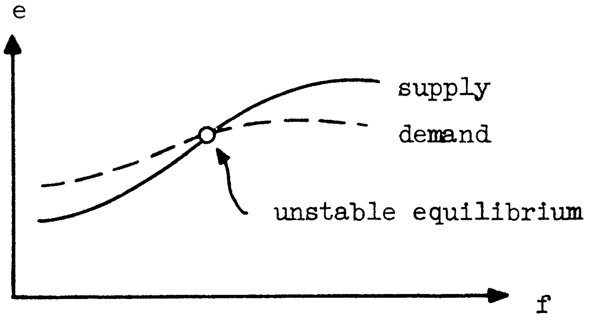
Thus far we have only spoken in general terms about the depiction of power bonds, presumably in preparation for a detailed quantitative analysis, and have not concerned ourselves specifically with what happens when two systems are coupled together. Naturally the power transferred across a bond is a function of the characteristics of the two participating systems. It is generally possible to conceive of one of the systems as the "supplier" and the other as the "recipient" of power. Consider two coupled systems, S_1 and S_2, operating in the steady-state:

$$S_1 \rightleftharpoons S_2$$

Now, it is plausible to assume that S_1 has a falling e-f characteristic such that e decreases as f increases. It is also to be expected that S_2 has a rising characteristic such that e and f increase together. If this happens to be the case the steady-state equilibrium point will correspond to the intersection of the two characteristics.



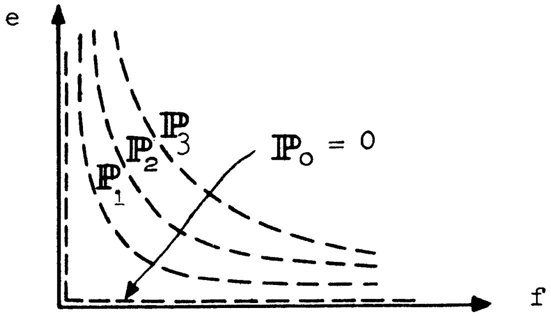
If the e-f characteristics of S_1 and S_2 are such that there is no definite intersection, then there exists no point of equilibrium operation. On the other hand, there may be systems which possess characteristics as sketched below:



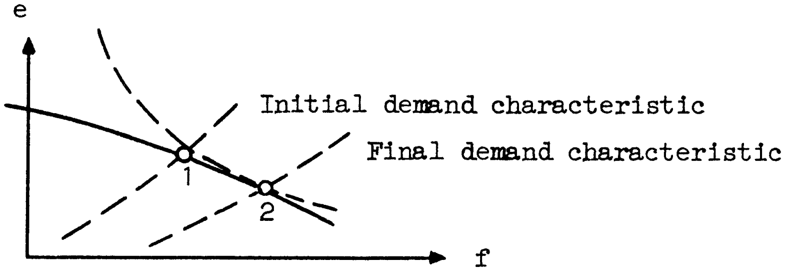
It is apparent that the intersection, though definite, generally corresponds to an unstable equilibrium point.

It is all very well to be able to determine the equilibrium power transfer, but we often wish to do much more than this. In particular, how can one maximize the power transferred in the steady-state operation of two coupled systems S_1 and S_2?

Since $\mathbb{P} = e \cdot f$, the curves of constant power are equilateral hyperbolas on the e-f plane.



Thus, plotting these curves on the same grid as the supply and demand characteristics of S_1 and S_2



it is apparent that the equilibrium point (1) is not optimal while (2) is. We may say in general that a design problem of a rather complex nature must be solved in order to match the characteristics of two coupled systems so as to achieve an optimum power transfer. The modification in S_2 which resulted in the movement of the equilibrium point from (1) to (2) is an example of "load matching." Presumably S_1 might have been altered instead to achieve a different optimal operating condition--an example of "source matching."

VI. Multiported Systems and Elements

A. Introduction

At this point it is clear that, once reticulated, any material system may be conceived as a multiported device with multiported elements. Thus, we shall consider briefly certain properties of multiports in general, and discuss in greater detail a particular universally encountered multiport, namely, the *ideal energy junction*.

> If the number and variety of multiport components of an engineering system is sufficiently large, more than one operable structure, circuit, or system could be assembled from the same parts. Abstractly, this is equivalent to the statement that for any given parts list (or *molecular* formula) more than one possible bond diagram (or *structural* formular) may exist. Systems which possess the same list of components but have differing bond structures may be considered as *structural isomers*, and the situation may be referred to as *structural isomerism*, borrowing a usage from chemistry. The number of possible structural isomers increases very rapidly as the number and variety of components increases. In this connection the equivalent situation in organic chemistry is instructive. For example, Butane [C_4H_{10}] has two isomers, Octane [C_8H_{18}] has eighteen, while calculations indicate that the homologous polymer $C_{40}H_{82}$ has 62,491,178,805,831 theoretically possible isomers! As to variety, more than one million diverse organic compounds have already been identified involving just the four atoms, C, H, O, and N, and this figure is growing exponentially with time. In the engineering systems field some aspects of structural isomerism have already been treated extensively in connection with circuit theory as we shall learn presently.

We have emphasized the necessity of abstracting from the many attributes of a system those properties that are essential to the delineation of the functional connectedness of its elements. Indeed, a truly incisive analysis is one which is detached from a specific material embodiment and which focuses upon the functions of the elements and the manner in which they are bonded together. The properties with which the elements of a reticulated system are endowed are transcendent properties, i.e., the artificial boundaries between hydraulics, electronics, and thermodynamics are largely overlooked. For example, a transducer is a two-port--an energy converter--and a concern as to whether the conversion is electromechanical, hydromechanical, or thermoelectrical is often secondary.

B. Multiports

Although the denumerably infinite universe of all possible combinations of one-, two-, and three-port elements is not sufficiently broad to enclose all conceivable energetic systems, its extent is so great as to include most systems of practical interest. Thus, we shall confine our attention primarily to one-, two-, and three-port elements and combinations thereof.

1. One-ports

A one-port may be thought of as a generalized impedance, some specific examples being resistance elements, capacitance elements, and inertance elements, together with all one-ported networks composed of such elements. The one-port is schematically represented in an energetic bond diagram simply as

A —

Thus, if we consider the universe of all one-port combinations, we note that it has but a single additional member, namely

A — B

2. Two-ports

A two-port may be conceived as a generalized transport process, i.e., a process by which energy is transformed, transmitted, or transduced. Thus, a communication system may be looked upon as a string of two-ports. The viewing of an ordinary triode amplifier as a two-port is generally accepted and is subject only to the assumption of a constant power supply, i.e., the power supply is located within the conceptual boundaries of the two-ported element. Hugh SKILLING in his text, Electrical Engineering Circuits, discusses the two-port, or two-port net as he calls it, from the standpoint of the electrical engineer. He schematically depicts various internal reticulations of the two-port and concerns himself primarily with the description of the transfer characteristics of such an element. Contemporary books on transistor theory and application, such as the text, Transistor Circuit Engineering, demonstrate that linear two-ports may be mathematically represented by way of a 2 x 2 transformation matrix.

The universe of all possible combinations of two-ports has but one member since the elements in the chain

$$— C — D — E — F —$$

may be coalesced to yield a single, equivalent two port

$$— G —$$

If we admit combinations of one- and two-ports, two new members are added, namely

$$A — G —$$

$$A — G — B$$

It is immediately evident, however, that the universe of one- and two-ports combinations is far too simple and restrictive.

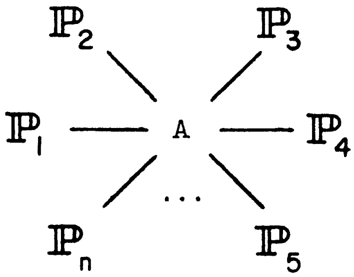
3. Three-ports

With the admission of the three-port the universe expands from the five members identified thus far to one which is denumerably infinite. The richness of this universe obtains from the possibility of branch structure which is attributable, of course, solely to the presence of the three-port. Thus, the three-port is a singular and most essential element.

We may think of the three-port as a generalized modulator, including (triportal) ideal energy junctions, power and signal modulators, power and signal amplifiers, and power exchangers as specific examples. Classical mechanics recognizes but a single three-port, namely the triportal energy junction; in this realm all systems are conceived as interconnected sets of one-ports (generalized impedances) and ideal energy junctions.

C. Ideal Energy Junctions

For a generic multiported element, A, the equation of energy continuity states:



$$-\sum_{i=1}^{n} \mathbb{P}_i = \frac{d\mathbb{E}}{dt} + \mathscr{P}_d$$

Now, let us restrict A to be ideal, by which we shall mean that $\mathscr{P}_d$ is identically zero, and if it further lacks the capacity for energy storage, then $\mathbb{E}$ vanishes as well, leaving simply the condition:

$$\sum_{i=1}^{n} \mathbb{P}_i = 0$$

A large class of energetic elements approximately satisfy this fundamental condition and the continued discussion of such elements is by no means trivial. Several of the most useful ideal elements are:

a. Energy Junctions
b. Ideal Transformers and Gyrators
c. Ideal Transducers
f. Differentials
g. Ideal Structural Modulators

In particular, we shall presently concern ourselves with the class of ideal energy junctions.

From this point onward a duality of characteristic relationships must be emphasized, for there exist two conjugate energy junctions--the effort junction and the flow junction. We may best depict this duality by carrying the development of both types in parallel, thus:

……

Effort Junction	Flow Junction

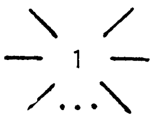
Both junctions are characterized by the condition that one of the two conjugate variables is common to all bonds, i.e., for a junction with n bonds:

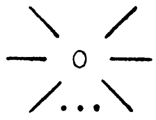
$f_i = f \quad (i = 1, 2, 3, \ldots, n)$	$e_i = e \quad (i = 1, 2, 3, \ldots, n)$

Then it immediately follows that

$\sum_{i=1}^{n} e_i = 0$	$\sum_{i=1}^{n} f_i = 0$

hereafter we shall represent the two junctions respectively as





……

The conjugate relationships

$f_i = f \quad (i = 1, 2, 3, \ldots, n)$ $\sum_{i=1}^{n} e_i = 0$ (Loop Law)	$e_i = e \quad (i = 1, 2, 3, \ldots, n)$ $\sum_{i=1}^{n} f_i = 0$ (Node Law)

play a dominant role in the idealized analysis of energetic systems. Carlo FERRARI depicts this role for a variety of media, perhaps the most familiar of which are the electrical network and the mechanical linkage. The conjugate junctions law are simple generalizations of KIRCHHOFF's Loop and Node Laws in the electrical case, and, borrowing FERRARI's terminology, the Laws of Velocity and Equilibrium in the mechanical case. The paper by J. C. SHOENFELD and the text by M. F. GARDNER and J. L. BARNES develop various aspects of the electromechanical analogy. In particular, SCHOENFELD noted that the flow junction (Node Law) in an electrical network and the effort junction (Equilibrium Law) in a mechanical system were isomorphic; thus, the importance of the duality in the concept of the general energy junction is underlined.

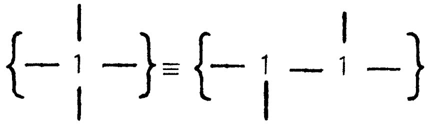
> The student, KIRCHHOFF, based upon a query in Neumann's physics seminar at Koenigsburg, made the first comprehensive study of the general electrical network problem by showing the relation between coarse reticulations (macroreticulations) and the field theorems (microreticulations). This was carried out in terms of Stoke's Theorem (Loop or Effort Conservation) and Gauss's Theorem (Node or Flow Conservation); the results were published first as an appendix to a paper in 1845 and then in more complete detail in 1847.
>
> As a warning against offhand use of the terms "Kirchhoff's First Law" and "Kirchhoff's Second Law," it is interesting to note that the laws appeared as follows in the two papers:

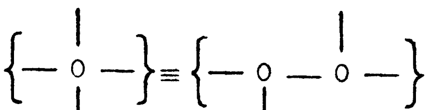
1845	1847
1) $I_1 + I_2 + \dots + I_\mu = 0$	I. $\omega_{k1} I_{k1} + \omega_{k2} I_{k2} + \dots = E_{k1} + E_{k2} + \dots$
2) $I_1 \cdot \omega_1 + I_2 \cdot \omega_2 \dots + I_\nu \omega_\nu =$ sum of the EMF	II. $I_{\lambda 1} + I_{\lambda 2} + \dots = 0$

Thus the node and loop rules are transposed in the two papers.

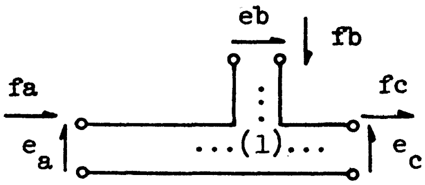
Implied in the assertion $e_i = e$ $(i = 1, 2, 3, \dots, n)$ and $f_i = f$ $(i = 1, 2, 3, \dots, n)$ is the assumed uniformity of the energetic medium by which the energy junction A is communicating with its environment. In other words, the bonds are either all electrical conductors or all mechanical links or all fluid-carrying ducts, etc.

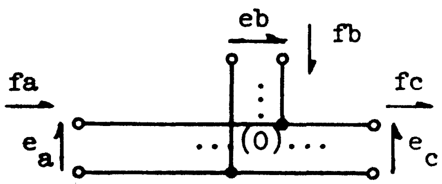
Energy junctions are associative and dissociative with respect to the triportal primitives; therefore, any multiported junction may be conceived as the combination of several three-ports. Thus for example:





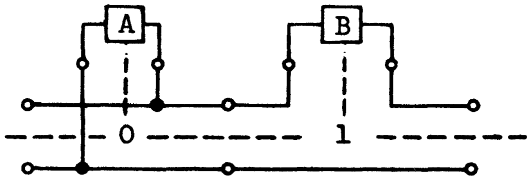
Historically, the notion of the energetic junction, in all its generality, has not been exploited effectively. In the analysis of electrical networks the concept has been developed more extensively than in other areas--notably by way of Kirchhoff's Laws. As a result, investigators in heat transfer, hydraulics, and so forth, have often resorted to the contrivance of electrical analogs. The sophistication of electrical schemata undoubtedly contributed to the attractiveness of this approach. However, we see now that such artifices are unwarranted in the light of the general formulation here presented. Just the same, it is illuminating to depict energy junctions as series and parallel electrical networks as sketched below:



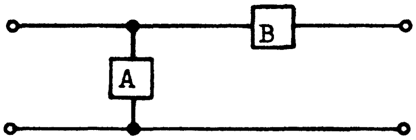


Let us next interpret the following interconnection of energy junctions and one-ports

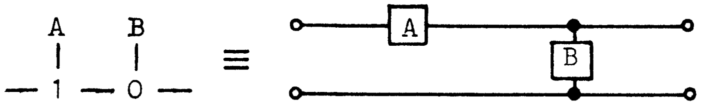
in terms of the equivalent electrical network. Recalling that A and B are generalized impedances, we have for the above



or, more simply



It is noteworthy that the circuit dual--the network resulting from a transposition $\left\{\begin{matrix} 0 \rightarrow 1 \\ 1 \rightarrow 0 \end{matrix}\right\}$ -- is immediately recognizable.



This, in fact, is an important attribute of a schematic representation in the form of the energetic bond diagram.

Background Reading

1. SKILLING, H. H. Electrical Engineering Circuits, Chapter 18.

   This reference gives some insight from the electrical engineering viewpoint into the nature and function of the two-port; attention is focused upon the problems of functionally or operationally describing such elements knowing their internal structure.

2. SHEA, R. F. (editor). Transistor Circuit Theory, pp. 1-3, 21-22, Appendix.

   The two-port is discussed relative to the description and analysis of transistors. One-ports and multiports are also mentioned. The matrix representation of linear elements is presented.

3. FERRARI, C. Relazione Generale sui "Modeli Analogie."

   This paper presents the conjugate junction laws for a number of important engineering media, thus lending breadth to the notion of the ideal energy junction.

4. SCHÖENFELD, J. C. Analogy of Hydraulic, Mechanical, Acoustic, and Electrical Systems.

   Recognition is given to the importance of the duality in the energy junction concept in delineating the electromechanical isomorphism.

5. GARDNER, M. F., and J. L. BARNES. Transients in Linear Systems, Chapter II.

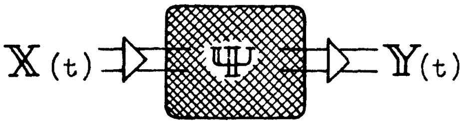
   Particular emphasis is given to the schematic representation of electrical and mechanical systems so as to exploit the analogies existing between them.

VII. Classes and Relations

A. Relations and Structure

Bertrand RUSSELL states: "To exhibit the structure of an object is to mention its parts and the ways in which they are interrelated." In the analysis of systems we are confronted with the task of establishing an order, a conceptual structure, in an initially formless universe. First, the S-E dichotomy is depicted in U, and then both S and E are further reticulated to the degree appropriate to the objectives of the analysis. The essential step in the process, however, is the recognition of the significant interrelationships among the reticulated elements. A system is not described by a parts list alone, but rather by the combination of a parts list and a delineation of the interconnections and interactions among the parts.

In dealing with relationships among the elements of a system, we must account for the existence of functional dependencies of the most general character. Thus far we have made mention of the generalized functional, Ψ, which scans the input vector $\mathbb{X}(t - \tau)$ for $0 < \tau < \infty$ and yields up a value for the output vector $\mathbb{Y}(t)$.



A particular form, and one for which we shall find frequent use in the sequel, is the vector-to-scalar transformation

$$Y(t) = \Psi[\mathbb{X}(t)]$$

For example, consider the correlation functional which yields the energy stored in an ideal element, namely:

$$Y(t) = \mathbb{E}(t) = \int_0^\infty \mathbb{P}(t - \tau)\, d\tau$$
$$= \int_0^\infty e(t - \tau) \cdot f(t - \tau)\, d\tau$$

Wherein we might identify $\mathbb{X} = \{x_1, x_2\} = \{e, f\}$.

Because of the fundamental role of relations in our analytical constructs we shall now turn to their general characterization in the context of the theory of classes, this being the most fundamental mathematical system available and therefore the most appropriate medium in which to couch a generalized description of relations.

B. The Concept of a Class

Out of a chaotic universe of sensory impressions and mental images, our reasoning mind struggles for order and understanding. The fundamental ordering principle upon which all this effort is based is that of likeness, resemblance or similarity. All thought springs from beginnings in comparative studies in which similar objects and phenomena are brought together into classes.

Thus the concept of a class (or alternatively, a set, collection, ensemble or aggregate) becomes the simplest component of mathematics and logical thought itself. The first step in establishing a class is that of determining the property of membership.

Membership

A class is determined (or established) the moment one arrives at a property (or rule, test or condition) which any object (or entity) within the universe under consideration must possess (or satisfy) in order to be a member of (or belong to) the class. Thus the concept of the class itself and the required rules for membership are inextricably interwoven. We shall inquire further into the nature of these conditions and properties below.

It is first worthwhile to introduce mathematical symbolism to make these concepts more precise. We shall accordingly denote various classes by Roman capitals:

CLASSES: A, B, C, etc.

The individual objects which comprise any of these classes we shall speak of as elements (or members or components) and denote by lower case Roman letters:

ELEMENTS: a, b, c, x, y, z, etc.

Properties possessed by the elements, including those properties upon which membership is based, will be denoted by small Greek letters:

PROPERTIES: α, β, γ, δ,, etc.

Between these elements and classes we have possible membership relations. The fact that a given element, a, is a member of a class, A, , we can conveniently express in the form

$$a \in A$$

by employing the membership symbol

MEMBERSHIP: $\in$: read "-- is a member of --"

Schematic diagrams frequently are used to aid in the comprehension of relationships between classes and elements. One approach is to depict the elements as geometric *points* and the classes as *sets* of points. Clearly, however, many other portrayals are possible. All such representations have justification to the degree that they lead to a self=evident or intuitive understanding of interrelationships.

It is frequently necessary to deny or negate the existence of a relationship between two objects; in a common symbolism the "operation" of *denial* is accomplished through the use of a vertical slash: " | ". Thus, to deny the membership relation we write, for example

$$b \notin A$$

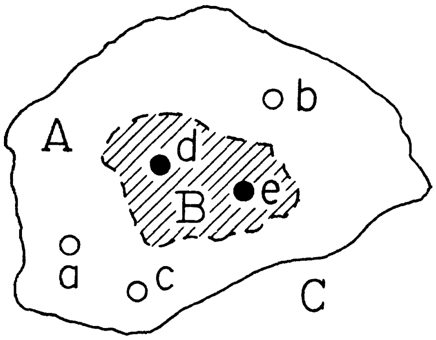
indicating that the element b does *not* satisfy the requirements for membership in the class A, that is, b is not a member of A.

No confusion should result if the relations of membership and nonmembership are stated in reverse fashion, thus:

$$A \ni a$$

$$A \not\ni b$$

Indeed, in the sequel, great significance will be attached to the so-called *converse relationships* of which those just above are examples.



Referring to the sketch we see that the following statements hold true:

$$a, b, c \in A \quad \text{but} \quad a, b, c \notin B$$

$$d, e \in B \quad \text{but} \quad d, e \notin A$$

$$a, b, c, d, e \in C$$

In an allegory we may liken the establishment of a class to the action of a small boy at the beach becoming interested in gathering white pebbles. We observe him gathering pebbles one by one, looking at them to see if they are white and either putting them into his pocket or throwing them back onto the beach. The defining property involved here is that of whiteness. All the pebbles in his pocket then are members of an evolving class of white pebbles and those thrown back belong to many other classes but in particular to the class of non-white pebbles. Any given pebble in his pocket can be considered as an element of the given class of white pebbles.

We may distinguish here at least two classes: the class A of white pebbles and the complementary class B of non-white pebbles. Any given pebble in the pocket we may distinguish by the lower case letter x and write the fact of membership in the form

$$x \in A.$$

Containment

Since most classification and gathering processes are not at any given time exhaustive, we must consider the existence of subclasses and the situation of containment within a class.

For example, in the allegory, the pebbles in the boy's pocket form a subclass *C* of the class of white pebbles *A*. By this we mean that *C* is part of but not all of *A*. We may symbolize this fact by the statement

$$C \subset A$$

employing the symbol

PART SYMBOL: $\subset$, read as "--is a (proper) part of --".

We refer to the class C as a proper subclass or *part* of the class Often, however, we do not wish, (or are not able), to establish the fact tha C is only a part of A, but wish merely to express the fact that C is contained in or included in A. We may still refer to C as a *subclass* of A but we may also wish to cover the possibility that C and A may be *coincident* or coextensive; that is, that C might include in some cases, each and every element of A. Thus a subclass is either a *part* or the *whole*. In this more general case we would write symbolically

$$C \subseteq A$$

CONTAINMENT: $\subseteq$ read as "--is $\left\{\begin{matrix}\text{contained}\\ \text{or}\\ \text{included}\end{matrix}\right\}$ in--"

It is interesting to note that the relations $\subset$ and $\subseteq$, which can hold between two classes, are respectively analogous to the relations $<$ and $\leq$ which may exist between two real variables. Indeed, this analogy can be a useful mnemonic device for those who are unfamiliar with the containment relations.

The operation of denial or negation may be applied to each of the containment relations, employing as before the vertical slash. Thus, for example,

$$D \not\subset A$$

As in the case of membership, it may frequently be convenient to write the containment relations in converse form. For instance, in the statement

$$A \supset C$$

the symbol, $\supset$, reads "--partly consists of--".

Identity of Elements

Two elements, a and b, of a given class are said to be equal or identical,

$$a = b$$

if they can be regarded as interchangeable with respect to the class and the associated class property. Thus, identity of elements only implies a certain relative indifference or indistinguishability within the context of a given class.

Elements which are not identical are said to be distinct and are indicated symbolically

$$a \neq b$$

Identity of Classes

In the pebble allegory, presume that through some quirk of geology, in a rather short time the boy had gathered together all the white pebbles on the beach. Then, if we continue to recognize the class C as the pebbles in his pocket and the class A as white pebbles, we might wish to express the fact that the class C had exhausted the class A; that is, each and every member of A was included in C. We can simply express the fact by the statement that the class C is identical to the class A, but we can also put very simple conditions on the two classes for this to be true.

Any two classes, X and Y, are said to be abstractly identical or equivalent, written:

$$X \equiv Y$$

f, and only if, $X \subseteq Y$ and $Y \subseteq X$. The equivalence symbol is read as follows:

QUIVALENCE: $\equiv$, read as "-- is $\left\{\begin{matrix}\text{equivalent}\\ \text{or}\\ \text{identical}\end{matrix}\right\}$ to --".

Thus the identity between the class of pebbles in the boy's pocket nd the class of white pebbles on the beach is established merely by deermining simultaneously whether all the white pebbles are in the boy's pocket nd whether all the pebbles in the boy's pocket are white pebbles.

Classes which have *no* elements in common are said to be *disjunct* or *disjoint*. We should particularly note that this is *not* the same as the fact that they are not equivalent, symbolized

$$X \not\equiv Y$$

which merely means that X and Y do not consist of the identical set of elements.

C. The Concept of a Relation

The purpose of the above discussion was to establish the context in which we shall seek an understanding of relationships or *relations* in their most general form. In doing this we inevitably encountered several specific relationships, namely those of *membership*, *containment*, and *identity* or equivalence. Each of these is an example of a relation between two objects or terms-- a so-called *diadic relation*, or simply a *diad*. The totality of objects linked by a given relation we call its *range* and it is thus apparent that a numbering of these objects affords a convenient approach to the classification of relations. That is, we may usefully distinguish between monads, diads, triads, tetrads, etc.

We shall here employ illuminated Roman capitals to denote relations, thus:

RELATIONS: $\mathbb{R}$, $\mathbb{T}$, $\mathbb{W}$, $\mathbb{X}$, etc.

If two relations, $\mathbb{R}$ and $\mathbb{S}$ are precisely the same, we may indicate their equivalence by way of the familiar notation

$$\mathbb{R} \equiv \mathbb{S}$$

The diad, "x bears the relation $\mathbb{R}$ to y" could be written symbolically either

$$x \, \mathbb{R} \, y \qquad \text{or} \qquad \mathbb{R}(x,y)$$

However, the nature of the terms or objects x and y is quite irrevelant: hence, the existential graph:

$$-\mathbb{R}-$$

with a specification of the realm of its applicability imparts the same information as the first form.

The <u>converse</u> of the diadic relation $\mathbb{R}$, when it has meaning, is written in our symbolism

$$- \reflectbox{$\mathbb{R}$} -$$

a form which is most suggestive of the significance of the converse. A particular form of the converse, namely the "inverse" of a mapping or transformation is often written $\mathbb{R}^{-1}$. It will shortly become evident that certain theorems which apply to transformations and their inverses also hold for relations when the inverse, $\mathbb{R}^{-1}$, is replaced by the more general converse, $\reflectbox{$\mathbb{R}$}$.

A restricted set of relations, namely those that express some form of identity, are <u>symmetric</u> in the terms such that $\mathbb{R} \equiv \reflectbox{$\mathbb{R}$}$. In writing such relations it is often convenient to employ a suggestive symbolism which exploits those letters that are inherently symmetric: $\mathbb{T}$, $\mathbb{H}$, $\mathbb{O}$, etc.

With this introduction to abstract relations, it is now propitious to focus our attention on certain specific types of relations of immediate present value. The objective of our study will be the establishment of a secure basis from which we may approach the relationships to be encountered in the generalized analysis of systems with increased understanding and insight.

We may organize this treatment on the method of categorization briefly introduced above, namely that founded on an enumeration of the objects linked by a given relation.

<u>Monads</u>

The statement: "there exists the object x" is an example of a monadic relation or <u>monad</u>--its range is the single entity x. In customary mathematical symbolism it is written

$$\exists x$$

The monad is so simple, and its statement and structure so succinct, that one is hard put to elaborate upon it. However, a consideration of the grammatical structure of the literal statement monad is perhaps illuminating. Let us, therefore, examine in greater detail the exemplary monadic statement, "there exists the object x." The converse form, "the object x exists," suggests the symbolism

$$x\ \mathbb{E}$$

wherein the existence relation is denoted by an illuminated capital in conformance with our notation. We note that the "subject" and "kernel" of the statement "x exists" have their counterparts in the symbolic statement. That is,

	Literal Statement	Symbolic Statement
Subject	the object x	x
Kernel	exists	$-\mathbb{E}$

The denial of existence is accomplished with the application of the vertical slash, thus:

$$\not\mathbb{E} \quad \text{or} \quad \nexists$$

When applied to x, this new monad would be read:

"The object x does not exist."

The generic monad is written simply $-\mathbb{R}$, together with a specification of its field of applicability. That is, when we say

$$x\,\mathbb{R} \quad \text{or} \quad \mathbb{R}(x)$$

we imply that $-\mathbb{R}$ may be meaningfully applied throughout the class X whose elements x have the certain common property characteristic of the class.

Diads

As a result of the utter simplicity of the monad its significance as a relation tends to elude the intuitive grasp which one has for higher order relations. The diad, then, is the simplest relation that has a immediate intuitive significance. The range of the diad consists of two objects, a and b for example. Symbolically, the diadic relationship may be expressed

$$a\,\mathbb{R}\,b$$

An alternative form may at times be appropriate; it is written

$$\mathbb{R}(a, b)$$

The first form has many mnemonic advantages and has by far the widest use, but we frequently employ both forms. It is noteworthy, in the second form, that in general the commutation of the terms inside the parenthesis is not valid. That is,

$$\mathbb{R}(a, b) \quad \not\equiv \quad \mathbb{R}(b, a)$$

We may represent a diad by an existential graph if to the relationship symbol $\mathbb{R}$ we append two tails thus

$$-\mathbb{R}-$$

which indicate its diadic nature. The relation $-\mathbb{R}-$ is denied or negated by application of the vertical slash, $-\not\mathbb{R}-$.

Three fundamental properties which are either present or absent in any diadic relation are the following:

1) Reflexivity: $a \mathbb{R} a$

   Any relation satisfying this condition is said to be reflexive; if $a \not\mathbb{R} a$, the relation is irreflexive.

2) Symmetry: If $a \mathbb{R} b$ then $b \mathbb{R} a$

   Any relation satisfying this condition is said to be symmetric; if $a \mathbb{R} b$ but $b \not\mathbb{R} a$ then the relation is asymmetric. A third important possibility is that of antisymmetry: $a \mathbb{R} b$ and $b \mathbb{R} a$ if and only if $a \equiv b$.

3) Transitivity: If $a \mathbb{R} b$ and $b \mathbb{R} c$ then $a \mathbb{R} c$

   Any relation satisfying this condition is said to be transitive; otherwise it is said to be intransitive.

It is possible to regard any diadic relation as directed or polarized. That is in an existential graph:

$$a \longrightarrow \mathbb{R} \longrightarrow b$$

Corresponding, then, to any such polarized relation, $\mathbb{R}$, there will often be a unique and well-defined converse, $\reflectbox{$\mathbb{R}$}$, such that if $a \mathbb{R} b$ then $b \reflectbox{$\mathbb{R}$} a$. It is important to note, however, that not every $\mathbb{R}$ possesses a meaningful converse. Referring to the definition of symmetry, it becomes obvious that relations which are symmetric not only do possess converses, but in addition satisfy the condition that

$$\mathbb{R} \equiv \reflectbox{$\mathbb{R}$}$$

In the light of the above discussion let us consider two all-important classes of diads--the abstract equivalence and ordering relations. These will play a dominant role in the development of the generalized functions and transformations which are required to describe the behavior of real systems.

Equivalence Relations. A diad, - $\mathbb{I}$ -, is said to be an equivalence relation if, and only if, $\mathbb{I}$ is: (i) reflexive; (ii) symmetric; and (iii) transitive.

So far we have encountered the two equivalence relations--Identity of Elements (=) and the Equivalence of Classes ($\equiv$) -- but there are many other examples of equivalence relations in all branches of mathematics and logic--in particular, Congruence ($\cong$) and Similarity ($\sim$) in ordinary Euclidean geometry.

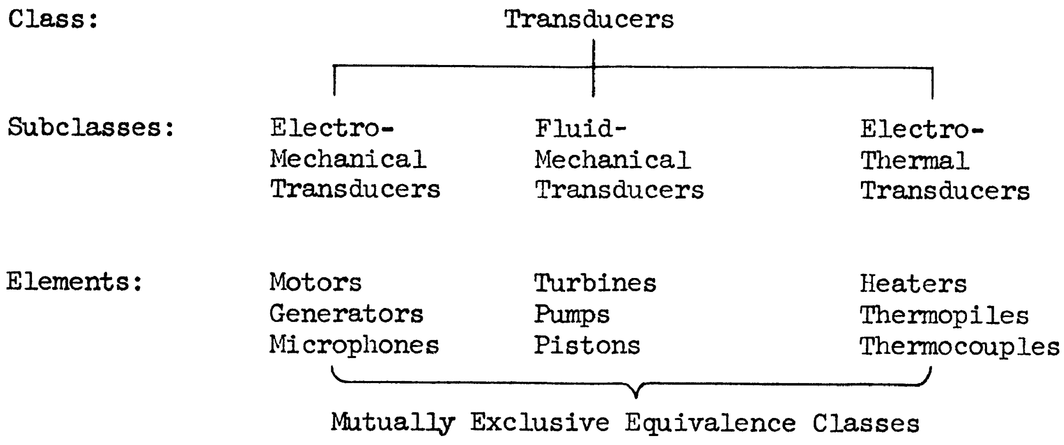
The fundamental property of any equivalence relation is that it divides the range over which it applies into a k-fold set of mutually exclusive equivalence classes ($\mathbb{K}^k$). Thus,

$$a \; \mathbb{I} \; b, \text{ if, and only if } (a, b) \equiv \mathbb{K}^k$$

The number, k, may be either finite or infinite, and in the latter case, either denumerable (i.e., countable) or nondenumerable.

In the simplest possible case the range will be merely bisected giving rise to a dichotomy or dichotomic categorization (e.g., up, down; positive, negative; etc.). Thus, such classification systems will in general give rise to polychotomies or manifold categories as indicated in the following table:

Example of Equivalence Classes



In all cases, any element in a subclass is abstractly equivalent or identical, *but only with respect to the defining property or condition* of the subclass. For example, it is obvious that the prime number 7 is not "equal" to the prime 11 in the sense of ordinary arithmetic, but only "equivalent" in the sense that they are both prime numbers.

Ordering Relations. If we are given any asymmetric ordering relation, - $\mathbb{J}$ -, applicable over a range (a, b) in a class $\mathbb{K}$, we can construct a corresponding antisymmetric ordering relation - $\mathbb{D}$ - by defining $\mathbb{D}$ to be the same as either $\mathbb{J}$ or $\mathbb{I}$ where — $\mathbb{I}$ — is an equivalence relation.

We may say then that $\mathbb{J}$ is a *strong* or *serial ordering relation*, it being: (i) *irreflexive*; (ii) *asymmetric*, and (iii) *transitive*. On the other hand, $\mathbb{D}$ is a *weak* or *partial ordering relation* since it is: (i) *reflexive*, (ii) *anti-symmetric*, and (iii) *transitive*.

The following table gives examples of these ordering relations which are already familiar to us:

Range	$\mathbb{J}$	$\mathbb{I}$	$\mathbb{D}$
Real Numbers	$>$	$=$	$\geq$
Classes	$\supset$	$\equiv$	$\supseteq$

It is important that we distinguish between the various ways diads may be combined. The three diad combinations which we shall briefly consider are: (i) *composition*, (ii) *alternation*, and (iii) *conjunction*.

Composition. Suppose x $\mathbb{R}_1$ y and y $\mathbb{R}_2$ z. Then x $\mathbb{R}_3$ z where

$$\mathbb{R}_3 = \mathbb{R}_1 \Diamond \mathbb{R}_2$$

the symbol $\mathbb{R}_1 \Diamond \mathbb{R}_2$ denoting the *composition* of the two relations $\mathbb{R}_1$ and $\mathbb{R}_2$. By way of an example, suppose $\mathbb{R}_1 \equiv \mathbb{M}$ and $\mathbb{R}_2 \equiv \mathbb{F}$ where $\mathbb{M}$ and $\mathbb{F}$ are respectively the relations of motherhood and fatherhood. Then, if $\mathbb{R}_3 = \mathbb{M} \Diamond \mathbb{F}$ and x $\mathbb{R}_3$z , then x is the *paternal grandmother* of z.

In connection with the composition of two relations it is noteworthy that if

$$\mathbb{P} \Diamond \mathbb{Q} \equiv \mathbb{R}$$

then

$$\reflectbox{$\mathbb{Q}$} \ \Diamond \ \reflectbox{$\mathbb{P}$} \equiv \reflectbox{$\mathbb{R}$}$$

which is seen to be a generalization of the more familiar statement for transformations:

$$\mathbb{Q}^{-1}\,\mathbb{P}^{-1} = (\mathbb{P}\,\mathbb{Q})^{-1}$$

Furthermore, we note that in the case of any transitive relation $\mathbb{T}$ that

$$\mathbb{T} \ \Diamond \ \mathbb{T} \equiv \mathbb{T}$$

Alternation. The alternation of two relations $\mathbb{R}_1$ and $\mathbb{R}_2$, written

$$\mathbb{R}_3 = \mathbb{R}_1 \cup \mathbb{R}_2$$

is the result of applying either $\mathbb{R}_1$ or $\mathbb{R}_2$. Thus, if again $\mathbb{R}_1 \equiv \mathbb{M}$ = motherhood and $\mathbb{R}_2 \equiv \mathbb{F}$ = fatherhood, then $\mathbb{R}_3 = \mathbb{M} \cup \mathbb{F}$ is the relation of parenthood.

Conjunction. The conjunction of two relations $\mathbb{R}_1$ and $\mathbb{R}_2$ written

$$\mathbb{R}_3 = \mathbb{R}_1 \cap \mathbb{R}_2$$

is the result of applying both $\mathbb{R}_1$ and $\mathbb{R}_2$. Thus, $\mathbb{R}_3 = \mathbb{M} \cap \mathbb{F}$ is, in mammalian biology, impossible since there is no x and y such that x $\mathbb{M}\cap\mathbb{F}$y, i.e., x is both the mother and the father of y. This fact may be expressed by asserting that $\mathbb{M} \pitchfork \mathbb{F}$ is true.

Triads

Any relation $\mathbb{R}$ whose range consists of three terms is a triadic relation or a triad, indicated in symbolic form

$$\mathbb{R}\,(a, b, c)$$

or existentially:

$$-\underset{|}{\mathbb{R}}-$$

A specific triadic relationship is, "b is between a and c." Moreover, any operation or rule of combination by which two terms "produce" a third term is a triad. A triadic relationship exists between a mother, a father, and a child, for example. In algebra and arithmetic, instead of saying two given numbers a and b determine a third number, c, such as their sum

$$c = a + b$$

or their product

$$c = ab$$

we may say that the three terms satisfy a triadic relation $\mathbb{R}$ among a, b, and c.

The structural or topological properties of triadic relationships are not so simple as those of diadic relations. However, several possibilities varying from complete symmetry to various types of asymmetry may be distinguished.

An example of complete symmetry occurs in the relationship between the three sides or the three vertices of an equilateral triangle. Here clearly

$$\begin{aligned} \mathbb{R}(a, b, c) &= \mathbb{R}(b, c, a) = \mathbb{R}(c, a, b) \\ &= \mathbb{R}(c, b, a) = \mathbb{R}(b, a, c) \\ &= \mathbb{R}(a, c, b) \end{aligned}$$

We may also speak of this as *permutative symmetry* since all permutations are allowable and are equivalent. Such symmetry also occurs for example in the algebraic equation

$$a + b + c = 0$$

However, some triads are symmetric only over a part of their range, such as the examples of *sum* and *product* mentioned above which are clearly symmetric in *a* and *b*; that is

Sum: $c = a + b = b + a$
Product: $c = a \, . \, b = b \, . \, a$

The "between" relation also possesses this limited symmetry since "*b* is between *c* and *a*"." It is appropriate that we designate such limited symmetry *commutative symmetry* by analogy to the commutative properties of the sum and product operators.

The relations of sum and product, as well as many other familiar triads, have an implicit polarization or directionality. That is, combining a and b *yields*, respectively, the sum, $c = a + b$, or the product, $c = a \, . \, b$;

the uniting of male and female results in the procreation of offspring so that the father-mother-child triad is also polarized. It is convenient to symbolize such directionality; for example,

$$\mathbb{R}\,(a,\ b;\ c)$$

or

$$\mathbb{R}((a,\ b) \rightarrow c)$$

or

$$c\,\mathbb{R}(a,\ b)$$

to indicate that c was the result of the combination of a and b.

Corresponding to certain types of asymmetric triads we can establish the existence of alternative relations. Consider, for example, the triadic relationship between father, mother, and child in which we distinguish at least the three polarized forms

$$\mathbb{R}_1\,(f,\ m;\ c)$$

$$\mathbb{R}_2\,(m,\ c;\ f)$$

$$\mathbb{R}_3\,(f,\ c;\ m)$$

Assuming normal wedlock these might be read as:

$\mathbb{R}_1$: c is the child of m and the child of f;

$\mathbb{R}_2$: f is the father of c and the husband of m;

$\mathbb{R}_3$: m is the mother of c and the wife of f.

We note that while $\mathbb{R}_1$ is symmetric in m and f, clearly $\mathbb{R}_2$ and $\mathbb{R}_3$ are asymmetric, but certainly in a broad sense the triadic relationship is established by each of the alternative forms so that they are, to this degree, equivalent.

A similar example to the above is the algebraic relation $x/y = z^2$

Polyads

With the monad, diad, and the triad we can build up a rich universe of polyadic relations--this because of the three-tailed property of the triad. That is, the relations

$$-\mathbb{M}$$

$$-\mathbb{D}-$$

$$-\underset{|}{\mathbb{T}}-$$

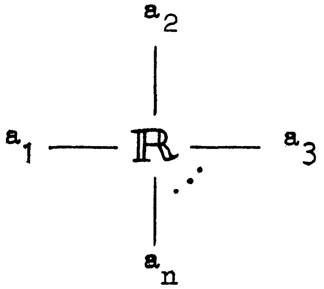
may be linked so as to construct a polyad of any order. However, the universe of polyadic relations so obtained is by no means exhaustive, although it is sufficiently broad for most of our purposes.

The abstract treatment of relations in mathematical literature has, for the most part, excluded the triad, concentrating instead on diads. It is self-evident, however, that the triad is essential if we are to consider even the simplest polyadic structures, since the compounding of diadic relations can never produce anything but diads. It is therefore possible to view the monad, diad, and the triad as the basic "building blocks" out of which all more complex relations can be constructed.

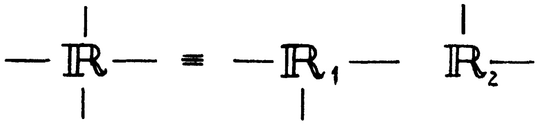
The general polyadic or "n-adic" relation may be symbolized

$$\mathbb{R}(a_1, a_2, \ldots, a_n)$$

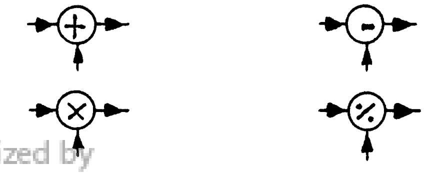
or, graphically



We may thus depict the reticulation of a tetrad into two triad primitives:



If $\mathbb{R}$ is an algebraic relation then it may always be reticulated into a system involving only the four triadic primitives:



However, for the generalized relations which are necessary to describe the behavior of real systems such a simple reticulation is not possible. Indeed, we shall see that there is a set of <u>logical triadic primitives</u> which is sufficiently general to serve as the building blocks for the construction of all such relations.

The relationships which bind together the characteristic variables of a physical system are clearly polyadic. We may think of them as falling into one or more of the following categories:

1) Correspondences
2) Functions
3) Transformations
4) Operators

We have been using the generic term <u>functional</u> to include functions, transformations and operators.

To properly describe the behavior of a system we must be ready to admit to the "functional domain" multivalued and discontinuous functions, unlike the traditional strategy of mathematicians, and indeed, many engineers. For example,

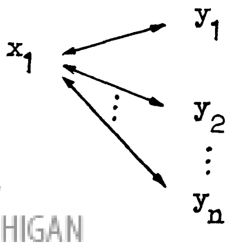
$$x = \sin^{-1} y$$

is a permissible relation. Certainly "a function" is a relation so that $y = F(x)$ is equivalent to the statement $y \; \mathbb{F} \; x$; the converse, $x = F^{-1}(y)$, if it is meaningful, is then written $x \; \reflectbox{$\mathbb{F}$} \; y$.

A correspondence is merely the statement that there is an x_1 which corresponds to a y_1, symbolized perhaps by the following notation:

$$\begin{array}{ccc} x_1 & \longleftrightarrow & y_1 \\ x_2 & \longleftrightarrow & y_2 \\ & \vdots & \\ x_n & \longleftrightarrow & y_n \end{array}$$

Multivaluedness may exist wherein the following correspondence could result; for example:



The generalized functional $\mathbb{Y} = \Psi(\mathbb{X})$ is one with which we are now familiar, and affords a sophisticated example of a polyadic relation.

Background Reading

(1) RUSSELL, Bertrand. Human Knowledge--Its Scope and Limits, Chapter III

This reference discusses aspects of structure as they pertain to the meaning of words, and the connexity of sentences and complex but meaningful sounds. It affords a valuable insight into the importance of relationships from a non-engineering viewpoint.

(2) PEIRCE, C. S. Collected Papers
See particularly
Vol. 3 The Logic of Relatives: 3.456 - 3.491
Vol. 4 Trichotomic Mathematics: 4.309 - 4.310
Vol. 5 The Valency of Concepts: 5.469

Without question, Peirce, the founder of pragmatism, was first to realize the singular character of the triadic relation. His use of bond diagrams for logical thought is prophetic and revealing. His philosophic concepts of Firstness (quality), Secondness (effect), and Thirdness (meaning) are grounded in the properties of monads, diads, triads, respectively. A word of caution -- Peirce's style runs the (deliberate?) gamut from extreme lucidity to perverse obscurity! But for those who like to climb mountains "just because they are there" Charles Sanders Peirce is a man to know.

(3) TARSKI, A. Introduction to Logic, Chapter V.

Tarski covers diadic relations in this text in a way which is easy to follow. He introduces the idea of a polyad, but without development.

(4) CHURCH, A. Introduction to Mathematical Logic, pp. 15-23.

The author discusses aspects of functions which are pertinent as background reading for the consideration of generalized relations.

(5) SUPPES, P. Introduction to Logic, Chapter 10.

The mathematical properties of (diadic) relations are discussed in an understandable fashion. Particular note should be given to the definition of anti-symmetry.

(6) BELL, E. T. Development of Mathematics, pp. 553-594.

This reference sketches the history of the development of mathematical logic from Leibniz (1666) until Godel (1931).

VIII. Continuum Logic and Hyperpolyhedral Functions

A. Introduction

A class of extremely flexible, n-dimensional piecewise-linear functions may be generated through the use of an extension of the logical operations of union and intersection. These hyperpolyhedral functions, as we shall call them, will be employed in the description and modelling of the behavior of physical systems. Such functions were first described by George Arthur PHILBRICK.

The union and intersection operations on classes have their basis in, and may be derived from, the diadic ordering relation, $\supseteq$. Thus, the process of comparison and subsequent establishment of order is fundamental to the development of the hyperpolyhedral functions.

B. Classes

Taking as fundamental the relations of membership and inclusion, $\in$ and $\supseteq$ respectively, and their denial, denoted by the vertical slash, $|$, the operations of logical union and intersection may be developed. If, for a given class X, the element $x \notin X$, then we define the complementary class X^* such that $x \in X^*$.

Union--Outer Selection

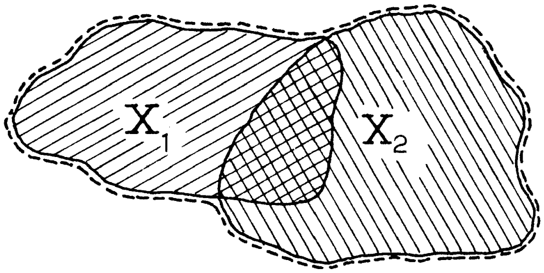
We desire a "least outer bounding class" $\underset{\smile}{X}$ which, for the aggregate $\{X_k \mid k = 1, 2, 3, \ldots, n\}$, satisfies the conditions that: (i) $\{X_k\} \subseteq \underset{\smile}{X}$; (ii) for every class $Y \supseteq \{X_k\}, \underset{\smile}{X} \subseteq Y$. We then define the operator $\mathbb{U}$ such that

$$\underset{\smile}{X} \equiv \overset{n}{\underset{k=1}{\mathbb{U}}} X_k \equiv \mathbb{U}(X_k) \equiv X_1 \cup X_2 \cup \cdots \cup X_n$$

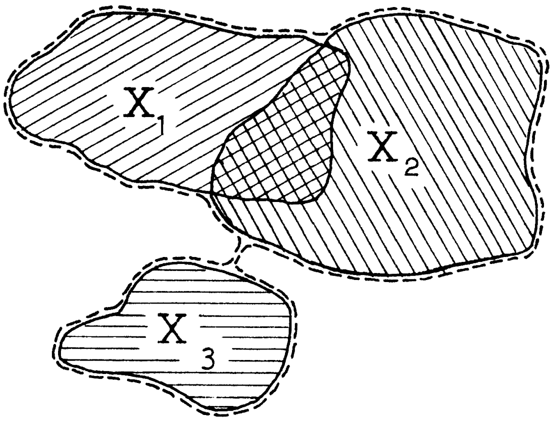
The class $\underset{\smile}{X}$, by definition, contains all the X_k, and, what is more, it is absolutely the smallest, most restricted class which does so. Thus, we refer to it as the least outer bound (l.o.b.) by analogy to real number theory. For two classes, X_1 and X_2,

$$\underset{\smile}{X} \equiv X_1 \cup X_2$$

wherein the symbol $\cup$ is often read "cup". The union operator for this case may be illustrated by depicting the classes as shaded areas; then, the union is enclosed by the dotted envelope:



If a third disjunct class X_3 is added, the union is as sketched below:



The union operation is associative and commutative; that is:

Associative property: $X_1 \cup (X_2 \cup X_3) \equiv (X_1 \cup X_2) \cup X_3$

$\equiv X_1 \cup X_2 \cup X_3$

Commutative property: $X_1 \cup X_2 \equiv X_2 \cup X_1$

If the aggregate $\left\{X_k \middle| k = 1, 2, 3, \ldots, n\right\}$ is extended without limit it becomes convenient to speak in terms of the universal class $\mathscr{S}$ such that, for any $\underset{\smile}{X}$ whatsoever,

$$\underset{\smile}{X} \subseteq \mathscr{S}$$

Intersection--Inner Selection

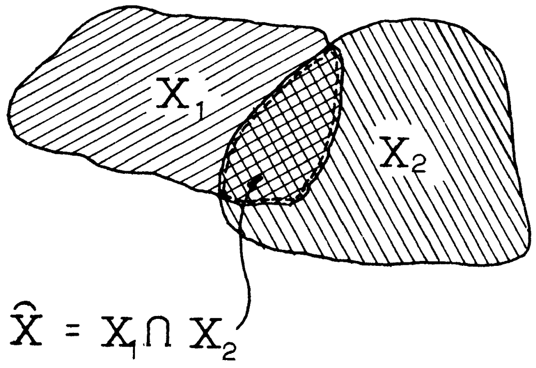
We now seek a "greatest inner bounding class" $\widehat{X}$, which, for the aggregate $\left\{ X_k \mid k = 1, 2, 3, \ldots, n \right\}$ satisfies the conditions: (i) $\widehat{X} \subseteq X_k$; (ii) for any class $Y \subseteq \left\{X_k\right\}$, $Y \subseteq \widehat{X}$. We then define the operator $\Cap$ such that

$$X \equiv \overset{n}{\underset{k=1}{\Cap}} X_k \equiv \Cap(X_k) \equiv X_1 \cap X_2 \cap \cdots \cap X_n$$

The class $\widehat{X}$ is, by definition, contained by all the X_k, and what is more, it is the largest, most extensive class which is so contained. Thus, we refer to it as the greatest inner bound (g. i. b.) by analogy to the greatest lower bound (g. l. b.) of real number theory. For two classes X_1 and X_2,

$$\widehat{X} = X_1 \cap X_2$$

wherein the symbol $\cap$ is often read "cap". The intersection operation for this simple case is illustrated by the sketch below:



The intersection operation is associative and commutative; that is:

Associative property: $X_1 \cap (X_2 \cap X_3) \equiv (X_1 \cap X_2) \cap X_3$

$\equiv X_1 \cap X_2 \cap X_3$

Commutative property: $X_1 \cap X_2 \equiv X_2 \cap X_1$

In expressions which involve both union and intersection these operations are mutually distributive. For example,

$$\begin{aligned}\bigcup[\bigcap(X_1,X_2,),\ X_3\ ] &\equiv X_3 \cup [X_1 \cap X_2] \\ &\equiv [X_3 \cup X_1] \cap [X_3 \cup X_2] \\ &\equiv [X_1 \cup X_3] \cap [X_2 \cup X_3]\end{aligned}$$

Obviously, if the aggregate $\left\{X_k \mid k = 1, 2, 3, \ldots, n\right\}$ includes one or more disjunctive classes then the class $\overset{\frown}{X}$ is empty, i.e., there are no elements $x \in \overset{\frown}{X}$. The concept of the empty or <u>null</u> class is an essential one; we denote this class by the symbol $\odot$. Evidently then,

$$\odot \subseteq \overset{\frown}{X}$$

Indeed, the following succession of inclusion relations holds for the classes we have defined thus far:

$$\odot \subseteq \overset{\frown}{X} \subseteq X_k \subseteq \underset{\smile}{X} \subseteq \$$$

Thus, if the aggregate $\left\{X_k \mid k = 1, 2, 3, \ldots, n\right\}$ includes all possible classes in the universe, $\$$, then, and only then, will

$$\overset{\frown}{X} = \overset{\frown}{X}_\infty = \overset{\infty}{\bigcap}(X_k) = \odot$$

$$\underset{\smile}{X} = \underset{\smile}{X}_\infty = \overset{\infty}{\bigcup}(X_k) = \$$$

<u>Complementation</u>

The concept of the <u>complementary</u> <u>class</u> is fundamental to the establishment of order. If, for a given class X the element $x \notin X$, then we define the complementary class X^* such that $x \in X^*$. As was done in the case of the union and intersection of classes, the <u>complementation</u> of a class X may be looked upon as an operation, the operator being denoted $(\quad)^*$.

C. Order

We have seen that the act of comparison among an aggregate of classes $\left\{X_k \mid k = 1, 2, 3, \ldots, n\right\}$, and the subsequent ordering by way of the diadic relation $\subseteq$, is basic to the establishment of the inner and outer bounding classes $\overset{\frown}{X}$ and $\underset{\smile}{X}$. Indeed, ordering is perhaps the most important operation in the universe; certainly it is fundamental to the establishment of any scale of measurement.

Lexicographic Order

The concept of order is often confounded with the idea of a scale of values or of numbers. It is important to demonstrate that ordering is independent of a number scale, and we shall do this by considering lexicographic order--the result of a weak ordering followed by additional weak ordering within the equivalence classes so produced, and continued until simple order is achieved. The prototype series, from which the name "lexicographic" is taken, may be thought of as an ordinary set of listings in a dictionary, telephone directory, or other lexicon.

In establishing such a series we first recognize that the universe $\$$, which is to be ordered, comprises the totality of letter groupings formed from the twenty-six letters and the blank space. A sample would be the following:

A
.
.
ABSOLUTE
ABSOLUTE ZERO
.
.
HYDROGEN
HYDROGEN ATOM
HYDROGENATE
HYSTERESIS
.
.
ZERO

Here, in establishing the series, the lexicographer first groups all entries into mutually exclusive classes based upon the initial letters (A, ..., H, ..., Z). Within each such weak ordered class there are in general several elements. Each of these elements is indifferent (or identical) with respect to any given equivalence class, say A. These may then be sorted into the proper subclasses A_, AA, ..., AH, ..., AZ. This sorting and arranging operation can be repeated until every element has been placed into a unique class such as the class ABSOLUTE. This ultimate class, consisting of but one element, constitutes a term in a series and will generally have a unique predecessor and successor, except in the case of the bounding classes A and ZZZ ...

Continuum Order

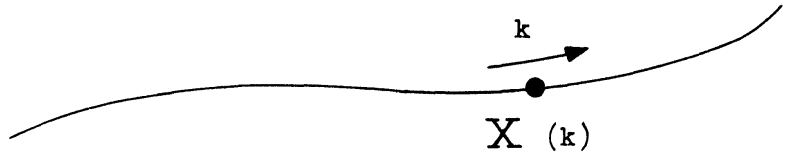
It is propitious at this point to consider an ordering wherein we suppose the universe $\$$ to envelope a single continuum of classes $\{X(k)\}$ Selecting any two classes from $\$$ we perform a test of comparison and establish which is greater in the abstract sense of the relation $\subseteq$; to the smaller of the two we assign $k = 1$, for instance, and to the larger $k = 2$. This process is repeated again and again over a large sample of the aggregate $\{X(k)\}$. The result (after possible renumbering) is an ordering such as

$$X(1) \subseteq X(2) \subseteq X(3) \subseteq \ldots \subseteq X(n)$$

If the process is extended ad infinitum, we can imagine that the continuum is completely ordered:

$$\odot \subseteq \overset{\frown}{X} \subseteq X(k) \subseteq \underset{\smile}{X} \subseteq \$$$

We wish now to define and to interpret the complementation of the ordered continuum. It is evident that the classes X (k) have been strung out along the ordered coordinate k in the fashion



Now if we were to convert or reverse the rank ordering; that is to rank from the "greatest" to "least", we would have

$$\$ \supseteq \underset{\smile}{X} \supseteq X(k) \supseteq \overset{\frown}{X} \supseteq \odot$$

We may now establish a biunique, 1-1 correspondence between the classes in the original order and those in the reverse order as follows:

$$\begin{array}{lccccccccccccccc}
\text{Original Order:} & \odot & \subseteq & \overset{\frown}{X} & \subseteq & X(1) & \subseteq & X(2)\cdots & & X(k_o) & & \cdots X(n) & \subseteq & \underset{\smile}{X} & \subseteq & \$ \\
\text{1-1 Correspondence:} & \updownarrow & & \updownarrow & & \updownarrow & & \updownarrow & & \updownarrow & & \updownarrow & & \updownarrow & & \updownarrow \\
\text{Converse Order:} & \$ & \supseteq & \underset{\smile}{X} & \supseteq & X(n) & \supseteq & X(n-1)\ldots & & X(k_o) & & \ldots X(1) & \supseteq & \overset{\frown}{X} & \supseteq & \odot
\end{array}$$

In general as a result of the homomorphism or biunicity the BROUWER theorem leads us to expect a fixed point, say k_o, which may or may not actually "exist" within the range of values. Moreover we can now define the class complementary to k as the mate in the correspondence above; then:

$$X^*(k) \equiv X(n + 1 - k)$$

Such converse order complementation may be conceived as a "reflection" or "rotation" of the continuum about the fixed point k_o. In physical measuring processes it is usually convenient (but not necessary) to take k_0 as 0 -- the physical zero or datum

It should be emphasized that neither a metric nor the concept of number is required to establish the ordered continuum; rather, order is founded on the simple act of comparison. Indeed, the establishment of order through comparison is a pre-requisite to the construction of a metric scale.

Upper and Lower Selection in the Continuum

Once order has been established in $\$$ then the operation of outer selection on the aggregate $\{X(k)\}$ will yield the class $\underset{\smile}{X}$ which corresponds, in the ordered scale, to the class $X(k_{max})$. That is,

$$\bigcup X(k) \equiv \underset{\smile}{X} \equiv X(k_{max})$$

In the same manner

$$\bigcap X(k) \equiv \overset{\frown}{X} \equiv X(k_{min})$$

Thus, in a sense, the operators $\Cup$ and $\Cap$ may be conceived as upper and lower selectors, respectively, in the ordered continuum. It is exactly in this sense that we wish to consider these operators in the discussion of continuum logic which follows.

D. Continuum Logic

The discussion up to this point has been completely general and unrestricted. It is now necessary to specialize to the case wherein the measure of the classes X(k) is some physical variable or value--perhaps, something as abstract and qualitative as a utility (as in the theory of games and decision-making), or something as concrete and quantitative as a weight, length, or voltage. We still need not suppose a scale--of weight, for example--to order the X(k) since the ordering could in this case be accomplished through the use of a balance. The establishment of a scale is a result, not a pre-requisite, of the ordering process.

With this specialization, the ordering relation $\subseteq$ particularizes to $\leq$ and the operators $\Cup$ and $\Cap$, which now are, in reality, upper and lower selectors, are to be interpreted as operators--special cases of generalized functional Ψ. According to its definition the operation $\Cup$ on an input bundle

$$\mathbf{X} \equiv \left\{ X\,(k) \right\}$$

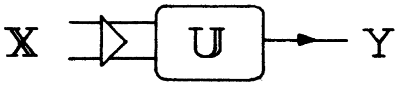
yields a single output value:

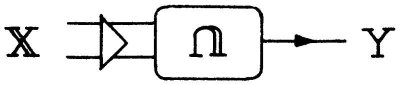
$$Y \equiv \Cup\,(\mathbf{X}) \equiv \underset{\smile}{X} \equiv \text{the greatest of the } X\,(k) \equiv X(k_{\max})$$

Similarly, in the case of the operator $\Cap$,

$$Y \equiv \Cap\,(\mathbf{X}) \equiv \overset{\frown}{X} \equiv \text{the least of the } X\,(k) = X\,(k_{\min})$$

There are many devices which can perform the operations $\Cup$ and $\Cap$, i.e., we can actually realize these operations in terms of computing elements as schematically depicted below:

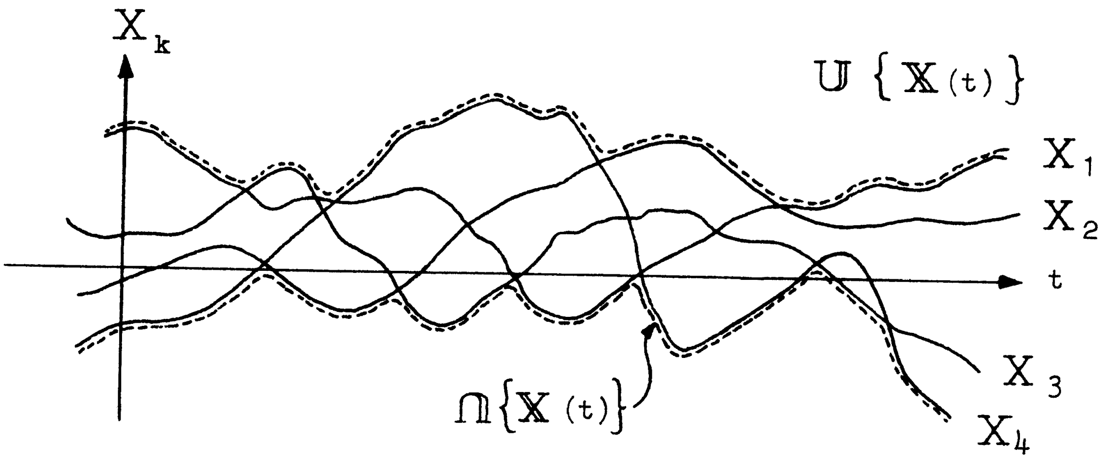




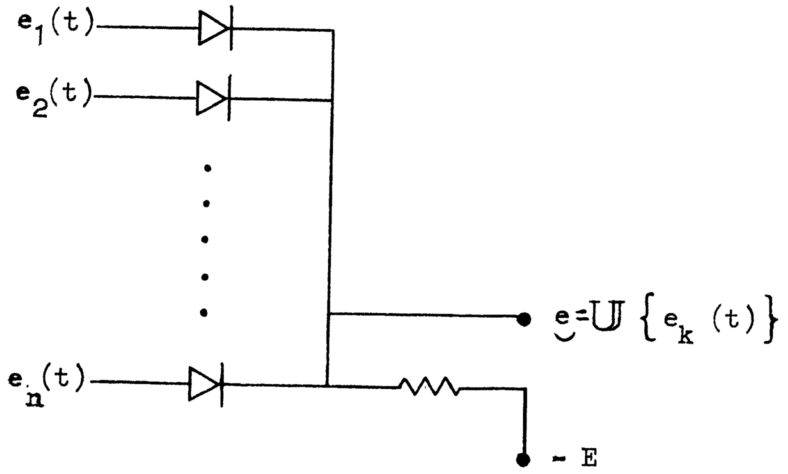
It is often the case that $\mathbb{X}$ is an ensemble of time-varying functions; that is,

$$\mathbb{X} \equiv \mathbb{X}(t) \equiv \left\{ X_k (t) \mid k = 1, 2, 3, \ldots, n \right\}$$

Then, of course $\Cup\{\mathbb{X}(t)\}$ will be the <u>greatest</u> of the X_k at time t, while $\Cap\{\mathbb{X}(t)\}$ will be the <u>least</u>. These then yield the dotted envelopes:

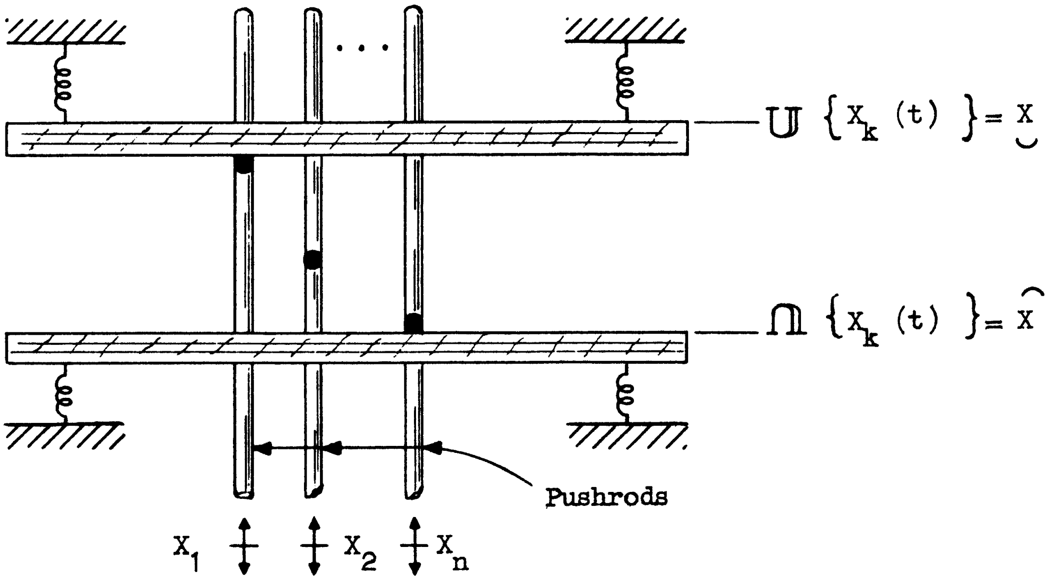


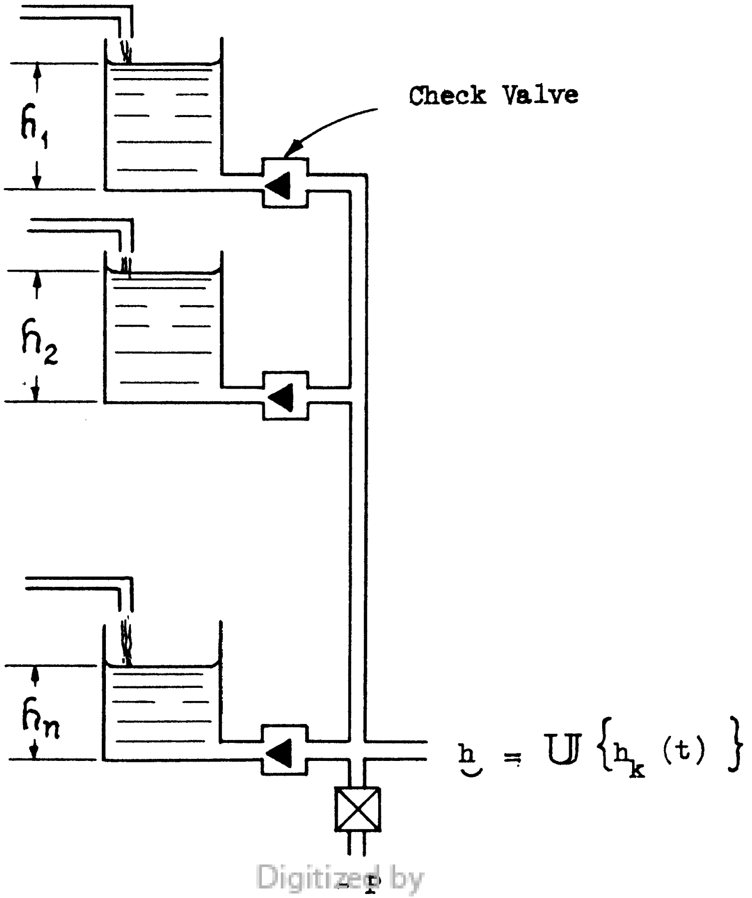
Perhaps the most common method for realizing upper and lower selection is the electrical scheme employing diodes. An upper selector is shown below:



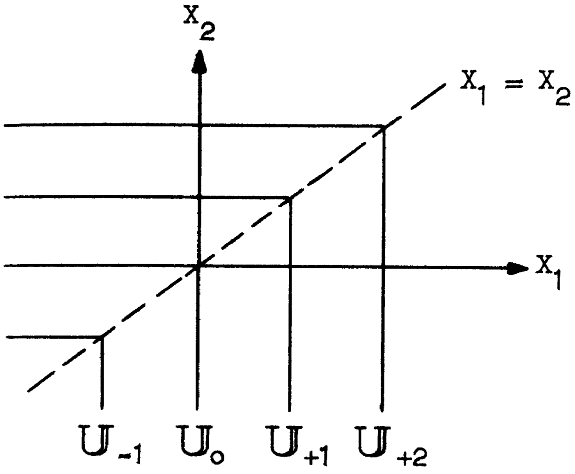
In electronics such a circuit has long been called an "auctioneer" circuit; in the same tenor the $\bigcap$ -circuit is called a "shopper".

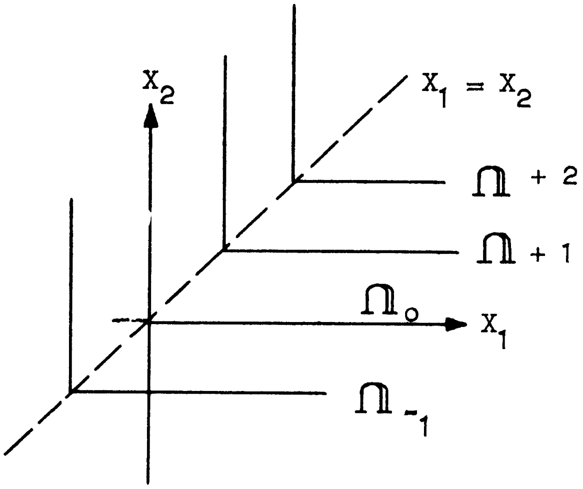
Of course, upper and lower selection may also be realized mechanically and hydraulically:





For the case where n = 2, i.e., where $\mathbf{X} = \{x_1, x_2\}$, it is informative to draw contours of the surfaces $\mathbf{U}(\mathbf{X})$ and $\mathbf{\cap}(\mathbf{X})$.





Thus, we see that there is no basic distinction between the operations $\mathbf{U}(x_1, x_2)$ and $\mathbf{\cap}(x_1, x_2)$, and, for example, the more familiar algebraic functions such as $\Phi_1(x_1, x_2) = x_1 + x_2$ or $\Phi_2(x_1, x_2) = x_1 . x_2$.

E. Two-Value or Binary Logic

As a particularization of continuum logic, wherein the X_k can assume any values whatever, we now consider two-value or binary logic wherein the X_k may be equal either to zero or to one. These two values are often taken as signifying, respectively, falsity (F) or truth (T). The convention is thus established that

$$0 < 1 \quad \text{or} \quad F < T$$
$$0^* \equiv 1 \quad \text{or} \quad F^* \equiv T$$

Note that the fixed point under complementary order reversal is non-existent (despite the fact we might call it "1/2")!

When there are two independent variables, i.e., when n = 2 in the aggregate $\{X_k \mid k = 1, 2, 3, \ldots, n\}$, the operations of union, intersection, and complementation may be conveniently portrayed in functional matrix form as shown below:

$$\begin{array}{c|c|c|c} & & X_1 & \\ & \cup(X_1,X_2) & 0 & 1 \\ \hline X_2 & 0 & 0 & 1 \\ & 1 & 1 & 1 \end{array}$$

$$\begin{array}{c|c|c|c} & & X_1 & \\ & \cap(X_1,X_2) & 0 & 1 \\ \hline X_2 & 0 & 0 & 0 \\ & 1 & 0 & 1 \end{array}$$

$$\begin{array}{c|c|c|c} & & X_1 & \\ & (X_1)^* & 0 & 1 \\ \hline X_2 & 0 & 1 & 0 \\ & 1 & 1 & 0 \end{array}$$

A second useful representation completely equivalent to the above is by way of the truth table, wherein the symbols T and F are substituted for 1 and 0.

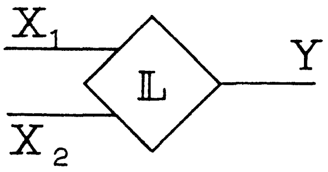
X_1	X_2	$\cup(X_1,X_2)$	$\cap(X_1,X_2)$	$(X_1)^*$	$(X_2)^*$
F	F	F	F	T	T
F	T	T	F	T	F
T	F	T	F	F	T
T	T	T	T	F	F

The following interpretations of the union, intersection, and complementation operations may now be made in classical binary logic:

Operation	Symbolic notation	Logical interpretation
Union	$\bigcup$	"( ) or ( )"
Intersection	$\bigcap$	"( ) and ( )"
Complementation	$(\quad)^*$	"not ( )"

There are a total of sixteen distinct binary operations $\mathbb{L}$, considering polarization; a convenient coding system may be employed as will be illustrated for the three operations introduced thus far. In the matrix form the elements are labelled generically $\alpha, \beta, \gamma, \delta$ as indicated below:

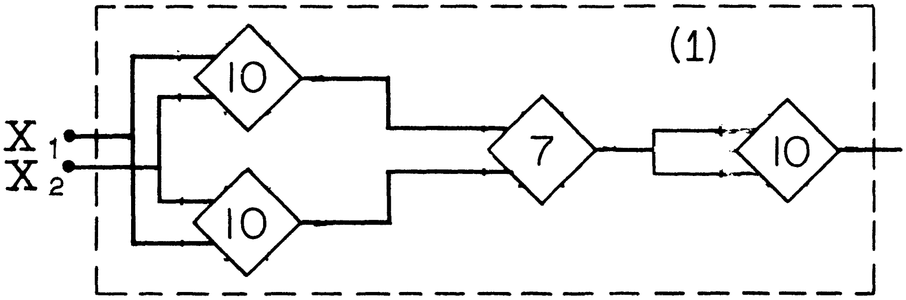
$$\begin{array}{cc|c|c|} & & X_1 & \\ & \mathbb{L} & 0 & 1 \\ \hline X_2 & 0 & \alpha & \beta \\ \hline & 1 & \gamma & \delta \\ \hline \end{array}$$



The entries for a given operation $\mathbb{L}$ are written in the fixed order $\alpha\beta\gamma\delta$, and the number for which these are the binary digits is then taken as the code number of the operation.

	Code	
Operation	Binary	Decimal
"or"	0111	7
"and"	0001	1
"not X_1"	1010	10
"not X_2"	1100	12

The sixteen binary operations are by no means independent. Indeed, they all may be established from the operations (1), (7), and (10). In fact, even this set of basic operations is not minimal for we may, for example, construct (1) from (7) and (10). The truth of this may be demonstrated by way of the signal flow graph shown below:



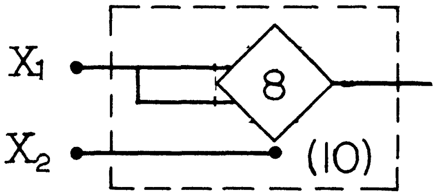
X_1	X_2	(10)	(10)	(7)	(10)
0	0	1	1	1	0
1	0	0	1	1	0
0	1	1	0	1	0
1	1	0	0	0	1

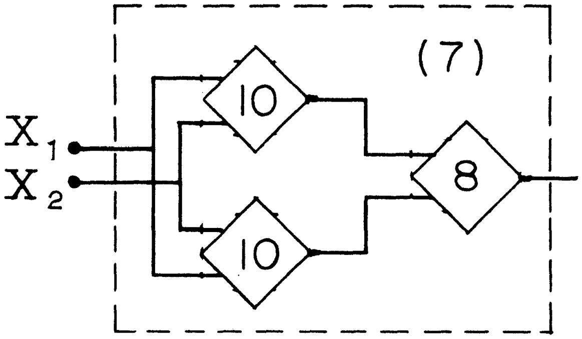
Indeed from the triplet (1, 7, 10) both the pair (7, 10) and the pair (1, 10) suffice as logical primitives, but the remaining pair (1, 7) does not.

The question now comes to mind: "Is there a single binary operation on the basis of which all sixteen operations may be established?" The two Sheffer Stroke operations "nor" (8) and "nand" (14) are each such complete logical primitives. A suggestive symbolism will be used for these operations, namely:

(8) ≡ ⊔↓⊔ ≡ nor ≡ not-or ≡ "dagger" (↓)
(14) ≡ ⊓|⊓ ≡ nand ≡ not-and ≡ "stroke" (|)

It is easy to demonstrate that these are indeed logical primitives; we shall simply verify that from (8) the operations (10) and (7)--which together are a set of logical primitives--may be constructed and we leave the remainder of the proof to the interested reader. The following signal flow graphs delineate the construction:





F. Multivalued Logic (Post Logics)

We may conceive of a spectrum of multi-valued or n-valued logics with binary and continuum logic occupying the extreme positions. If we think of the ends of the real line [o,1] as corresponding, respectively, to absolute falsity and absolute truth, then we might interpret all intermediate positions as corresponding to partial truth, as it were.

It is easy to construct the matrix form for the logical operations in such n-valued logics; an illustration is given below in the case of the operation U:

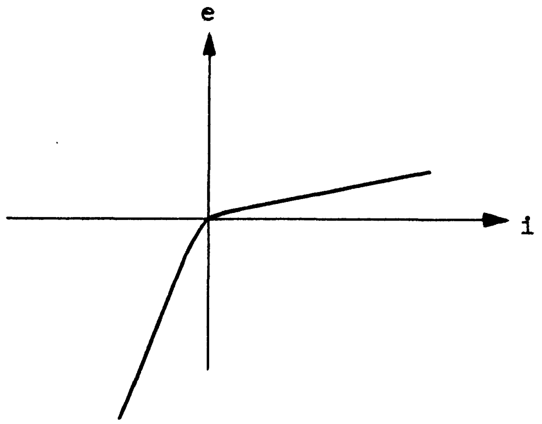
X_1

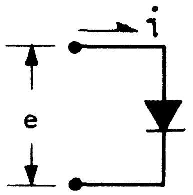
	$U(X_1,X_2)$	0	1	2	3	....	n
	0	0	1	2	3		n
	1	1	1	2	3		n
	2	2	2	2	3		n
X_2	3	3	3	3	3		n
	⋮						⋮
	n	n	n	n	n	....	n

The n-valued U and Π operators then give rise to Post algebras.

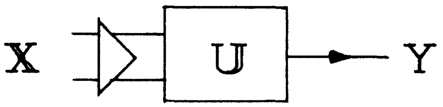
G. Hyperpolyhedral Functions

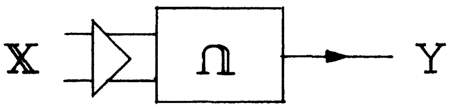
At the outset it was stated that a class of functions was sought which could be used to describe the behavior of physical systems in genera By proper choice of parameters such functions must be capable of conformance to not only continuous, well-behaved functional characteristics, but also to characteristics which are inherently nonlinear or discontinuous. Indeed, functions which play havoc with conventional mathematics must be rendered into articulate and tractable form in order to describe the behavior of many commonplace elements. Take, for example, the simple diode; plotting the voltage-current characteristic of a real diode generally yield a curve similar to the one sketched below:





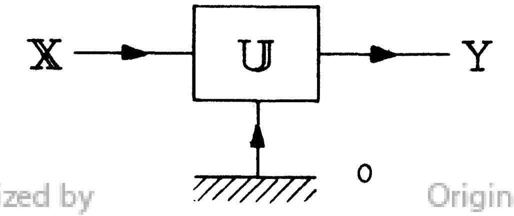
It is hardly necessary to point out the discontinuity in the slope of the characteristic at the origin. Thus, we seek a construct in the context of which all behavioral characteristics may be described. Such a construct is founded upon the operations of upper and lower selection, $\cup$ and $\cap$, i.e. the operators:



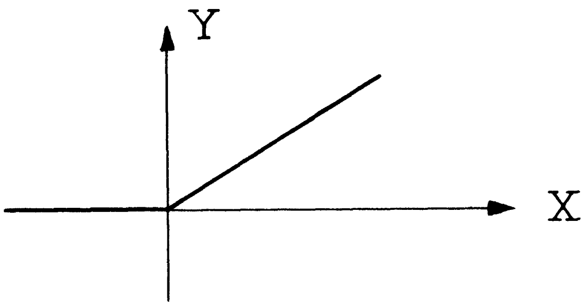


shall be utilized as the fundamental building blocks in the synthesis of complex, multi-dimensional functions.

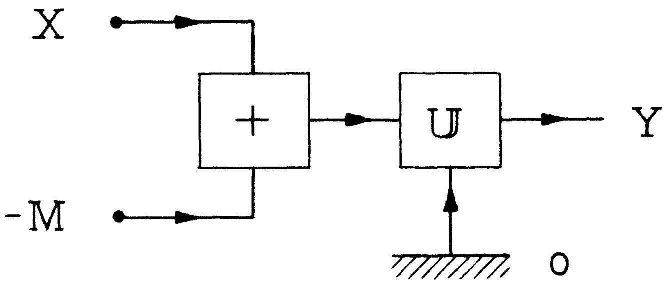
Consider, for example, the operation sketched below:

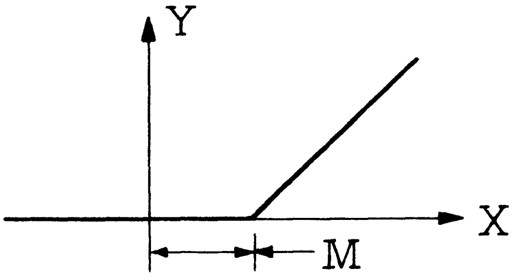


This yields a function $Y = Y(X)$ which may be sketched as follows:



If now we take as a unit cell or polygonal primitive



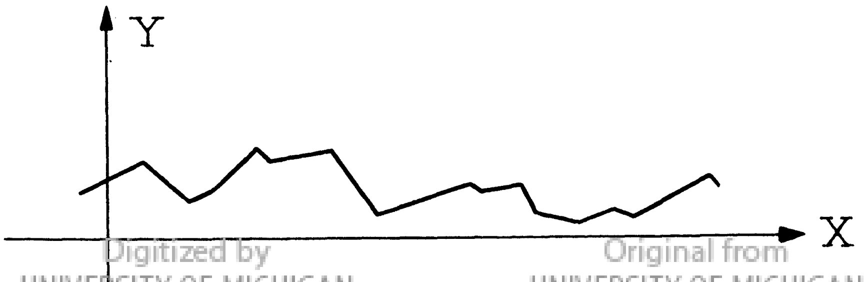


$$Y = U(o, X - M)$$

and then write

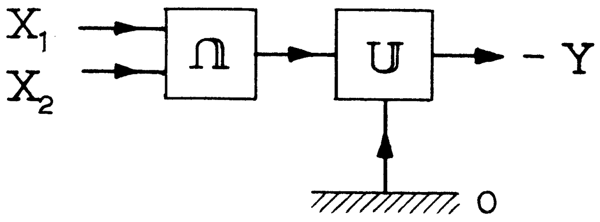
$$Y = \sum_{k=1}^{n} a_k \, U \, (o, X - M_k)$$

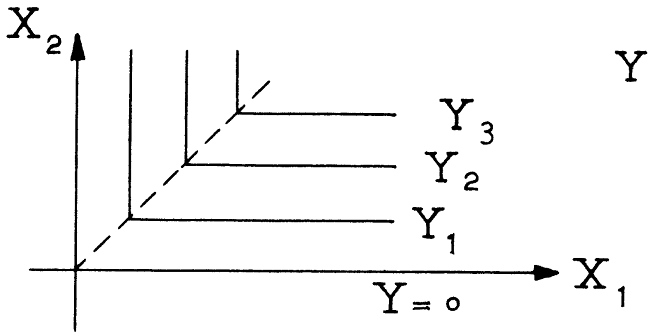
the result is a polygonal function which might appear as sketched below:



We can imagine that this is indeed a funicular polygon; i.e., a two dimensional curve consisting entirely of line segments.

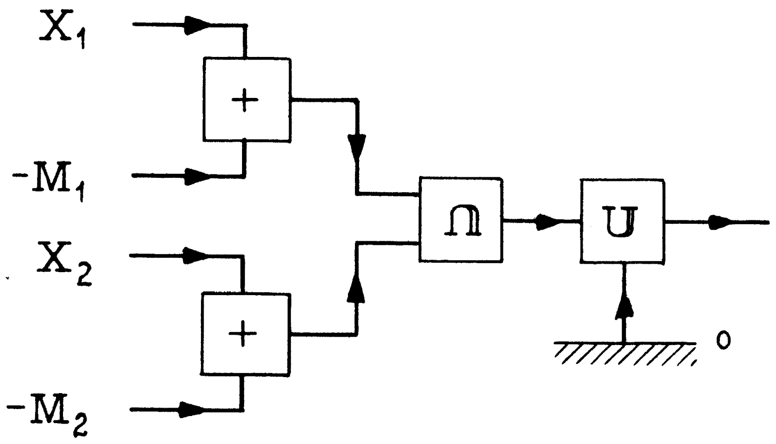
Next, we consider a more general operation which embodies both $\cup$ and $\cap$:

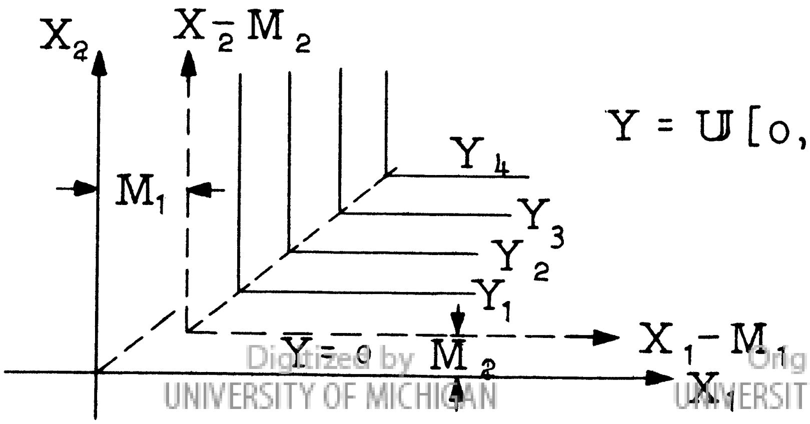




$$Y = \cup [o, \cap(X_1, X_2)]$$

An immediately evident generalization yields the polyhedral primitive,--



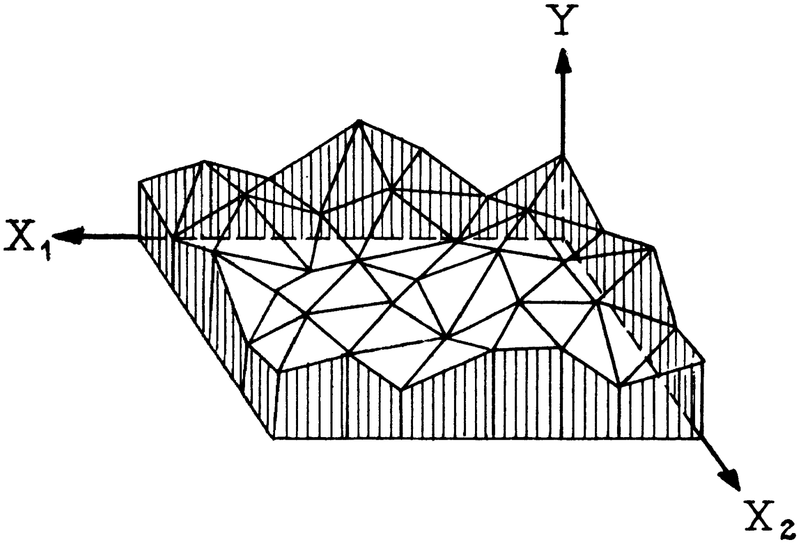


$$Y = \cup [o, \cap (X_1 - M_1, X_2 - M_2)$$

From which we construct the polyhedral function

$$Y = \sum_{j=1}^{n} \sum_{i=1}^{m} a_{ij} \mho [o, \Omega (X_1 - M_1^i, X_2 - M_2^j)]$$

A representation of this surface would reveal that it is composed entirely of triangular facets; the following sketch illustrates this:



Generalizing further, we arrive at the n-dimensional hyperpolyhedral function

$$Y = \sum_{i,j,k,\ldots u} a_{ijk\ldots u} \mho [o, \Omega (X_1 - M_1^i, X_2 - M_2^j, X_3 - M_3^k , \cdots X_n - M_n^u)]$$

$$= H_n \{ X_r \mid r = 1,2,3, \cdots n \} = H_n (\mathbb{X})$$

It is now a relatively simple conceptual step to hyperpolyhedral computing systems which utilize, as elements, hyperpolyhedral functions $H_n(X)$. We thus realize the fantastically variegated universe of functions at our disposal for modelling system behavior. It is to be emphasized, moreover, that such functions may be embodied in practical computing hardware.

Consider, for example, the polyhedral multiplier

$$Z = \sum_{i=0}^{n} \Cup [0, \Cap(X-i, Y)] + \sum_{j=1}^{m} \Cup [0, \Cap(X, Y-j)]$$

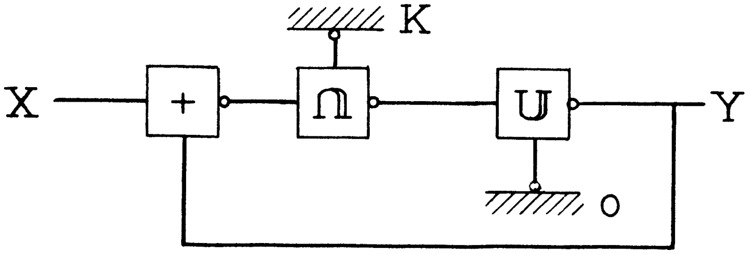
which is a first approximation to the product $Z = XY$. When X and Y are positive, this does, in fact, give exact results for integers.

Into the realm of hyperpolyhedral computing systems we certainly must admit implicit functions wherein the output depends upon itself as well as the inputs. That is,

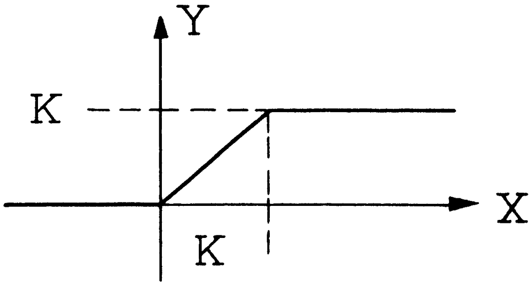
$$Y = \Phi(X, Y)$$

Schematically, the implicit feature appears as single or multiple feedback loops within the structure of the function which insert the output Y at various stages in the forward computation process. A simple example is the following:

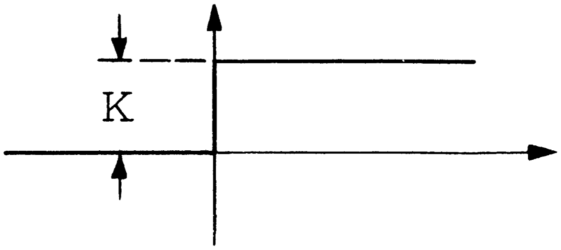
$$Y = \Cup [0, \Cap(K, X+Y)]$$



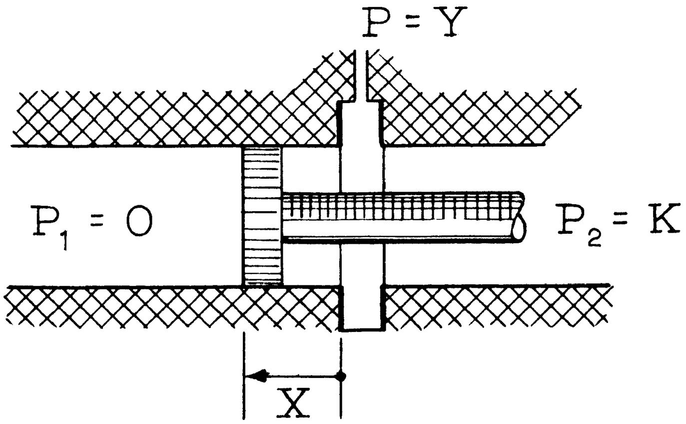
The open loop response characteristic is simply



With the addition of the feedback loop the response becomes:



Such a function might, for instance, be used to model the behavior of a control valve as sketched below:



Background Reading--Continuum Logic

1. Alonzo Church, Introduction to Mathematical Logic.

In the context of true-false or binary logic the important "connectives" or operators are introduced.

2. Paul Rosenbloom, The Elements of Mathematical Logic, pp. 51-65.

A discussion of multi-valued logic is given. The reader is cautioned to note that Rosenbloom interchanges the usage of $\cup$ and $\cap$.

3. Hans Reichenbach, The Theory of Probability, pp. 387-389.

The author discusses some of the implications of a multi-valued logic; the reader should note that n-valued logic occupies a position on a continuu at one end of which is binary logic and at the other end of which is continu logic.

Background Reading--Hyperpolyhedral Functions

1. George A. Philbrick, Continuous Electronic Representation of Nonlinear Functions of n Variables. (Palimpsest)

The author introduces the concept of piecewise linear functions built up from $\cup$, $\cap$, and " + " for use in analog computing when it is desired to fit an analog model to a body of empirical data.

2. Thomas E. Stern, Piecewise-Linear Network Analysis and Synthesis.

A formalism for dealing with piece-wise linear networks is developed from the fundamentals, although a rather unfamiliar nomenclature is used. Included are examples of polyhedral and pyramidal functions, as well as more sophisticated surfaces.

3. S. A. Ginsburg, Logical Method for Synthesizing Function Generators.

The development and viewpoint in this paper are basically similar to that in the Philbrick paper mentioned above. Here, however, somewhat more attention is given to the background of the logical constructs which give rise to the synthesis of piecewise linear functions.

4. G. A. Korn and T. M. Korn, Electronic Analog Computers. Chapter 6.

In the context of a general discussion of analog computer techniques are included brief descriptions of polyhedral multipliers and diode function generators.

5. H. J. Zimmerman and S. J. Mason, Electronic Circuit Theory.

Use is made of networks of ideal diodes in the synthesis of models of various essential circuit elements.

IX The Steady-State of Energetic Systems

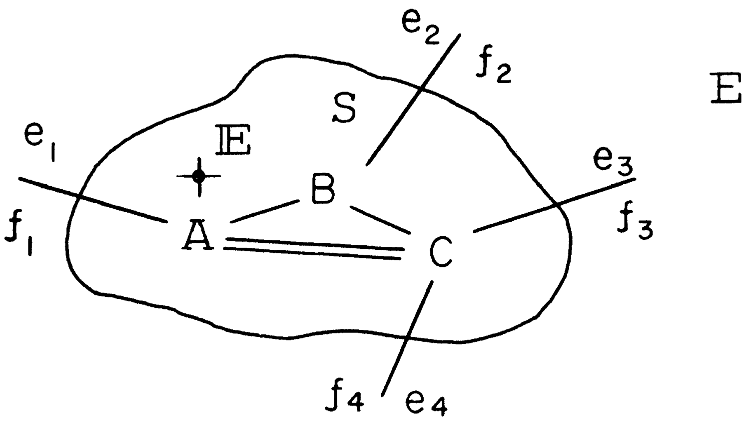
A. Introduction

The analysis of the steady-state plays a dominant role in providing us with an overall understanding of the behavior of energetic systems. Although the steady-state case, per se, is rather sterile and uninteresting (but by no means trivial!), an insight into the steady-state behavior of a system forms the basis upon which the analysis of its stability and transient behavior may be founded. Consider, for instance, the fact that the stability of a system may be evaluated by observing the result of small excursions about a steady operating point. Moreover, a transient condition in a stable system is the means by which its operating state alters from one steady condition to another.

We shall wish to distinguish between two types of steady-states: (i) the *static case*, wherein the power flux is identically zero and all that is required is a statement of the distribution of internally stored energy; (ii) the *stationary case*, wherein the power flux is *constant*, at least in the mean.

In rendering the description of the steady-state in terms of mathematical relationships we are faced with the problem of modelling all sorts of nonlinear, as well as linear, behavior.

B. The Static Case



Consider a four-ported system S. We define S to be static if it is in equilibrium with its environment E such that all the f's, both internal and external, are identically zero. If this is indeed the case, then certainly the power flow $\mathbb{P}$ is everywhere zero, and the state of the system is entirely described through a specification of the distribution of internally stored energy $\mathbb{E}$. In any real structure, which is inherently deformable, this is tantamount to a specification of the deformation of the system, i.e., the displacement of every particle thereof. Cases in point are electro- and hydro-static fields.

A static system, then, is one which has passed through some sort of transient condition during which power flows, from the environment and among the various parts of the system, were occurring. That is, we must conceive of the attainment of static equilibrium as a process requiring a finite interval of time. As the equilibrium state is approached, the power flows all decrease, and of course, vanish utterly when that state is reached. However, in any such process of practical interest there has been a net influx of power, leaving the system with internally stored energy which is distributed in a manner characteristic of the conditions on its boundary. The fact that such an energy distribution is ultimately reducible to a distribution of deformation leads us to consider displacement quantities

$$q = \int_{-\infty}^{t} f \, dt$$

In particular, at each of the ports we are concerned with the pair of conjugate variables (e_i, q_i), similarly, a pair (e, q) may be identified at each internal bond.

C. The Stationary Case

Referring to the sketch at the beginning of the previous section, if S is operating in a <u>stationary state</u>, then it is in dynamic equilibrium with E. In the case of <u>strict stationarity</u> the time derivatives of the f's both internal and external, vanish identically. A weaker condition is that of <u>quasi-stationarity</u> wherein the time averages of the $\dot{f}$'s are zero, i.e., each of the flows is fluctuating about some steady mean value.

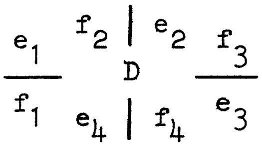
Strict Stationarity: $\dot{f}\ (\dot{t}) \equiv o$ or $f\ (t) = F =$ constant
for all bonds
$\therefore \quad \mathbb{P} \quad =$ constant
for all bonds

Quasi-Stationarity: $\overline{\dot{f}\ (t)} = o$ or $\overline{f\ (t)} = F =$ constant
for all bonds, e.g.
$f\ (t) = F_1 + F_2 \sin \omega t$
$\therefore \quad \overline{\mathbb{P}} \quad =$ constant
for all bonds.

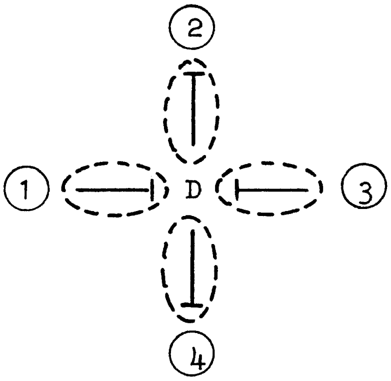
D. Determination of the Steady-State

By "determination of the steady-state" is meant an analysis which leads to a computation of the equilibrium magnitudes of the significant dependent variables of a system corresponding to a given set of independent variables, i.e., the impressed conditions at the boundaries. Such an analysis is by no means trivial in the case of a complex system.

It has been stated repeatedly that in order to perform any sort of incisive quantitative analysis the non-causal energy bond reticulation must be transformed into a causal bond reticulation. Thus, in the case of the four-port, D,



the assignment of causality might result in the diagram



Now, if there be any determinant stationary condition we must be able to write

$$f_1 = \Phi_f^1 (e_1, f_2, e_3, f_4)$$
$$e_2 = \Phi_e^2 (e_1, f_2, e_3, f_4)$$
$$f_3 = \Phi_f^3 (e_1, f_2, e_3, f_4)$$
$$e_4 = \Phi_e^4 (e_1, f_2, e_3, f_4)$$

If we are searching for a static condition then, of course, we replace all the f's by displacement quantities (q); hence, there must exist a set

$$q_1 = \Phi_q^1 (e_1, q_2, e_3, q_4)$$
$$e_2 = \Phi_e^2 (e_1, q_2, e_3, q_4)$$
$$q_3 = \Phi_q^3 (e_1, q_2, e_3, q_4)$$
$$e_4 = \Phi_e^4 (e_1, q_2, e_3, q_4)$$

In the above it is to be understood that the functions Φ are of the most general type. However, an important special case is that of linear functions, although in the real world true linearity is never found. Still, linear or <u>linearized</u> analysis facilitates the computational process and permits approximate answers to be obtained quickly. These are only in error to the degree that the system cannot be made to conform to a linear characteristic within its operating range. The fact remains, however, that one can quickly cite examples of elements which are essentially nonlinear in character (i.e., linearization is not possible); indeed, such nonlinearity is exploited by the designer and therefore cannot be overlooked by the analyst.

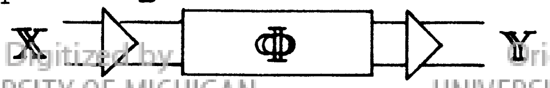
Suppose that Φ_f^1 is indeed linear; then, for small changes in the independent quantities,

$$\Delta f_1 = \mathcal{X}_1 \Delta e_1 + \mathcal{X}_2 \Delta f_2 + \mathcal{X}_3 \Delta e_3 + \mathcal{X}_4 \Delta f_4$$

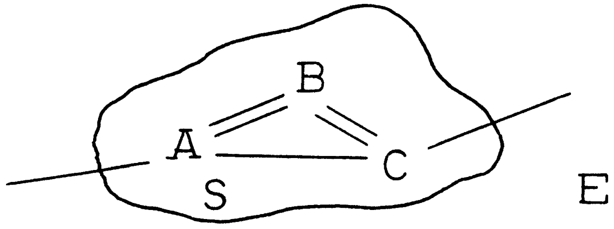
This, evidently, has reduced the problem to the ultimate in simplicity, yet the $\mathcal{X}_i$ ---the <u>influence coefficients</u> -- are not always easy to evaluate.

E. System Reticulation for Steady-State Behavior

The most general case of a static functional transformation, i.e., one which yields up an output value $\mathbb{Y}(t)$ corresponding to an input value $\mathbb{X}(t)$, is the operator $\mathbf{\Phi}$



In the present context we wish to consider practical measures which will facilitate the determination of the steady-state behavior, and to this end it is necessary that we reticulate the function Φ . Consider, for example, the two port

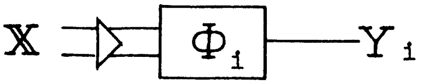


Corresponding to each bond there is a pair of conjugate variables, say e and f. Thus associated with this system are a total of fourteen quantities, only two of which are environmentally determined, i.e., are independent or <u>input</u> <u>quantities</u>. Hence there are twelve dependent or <u>output</u> quantities, each of which must be evaluated in order to specify the steady-state behavior. That is,

$$\mathbb{X} = \{ X_1 , X_2 \}$$

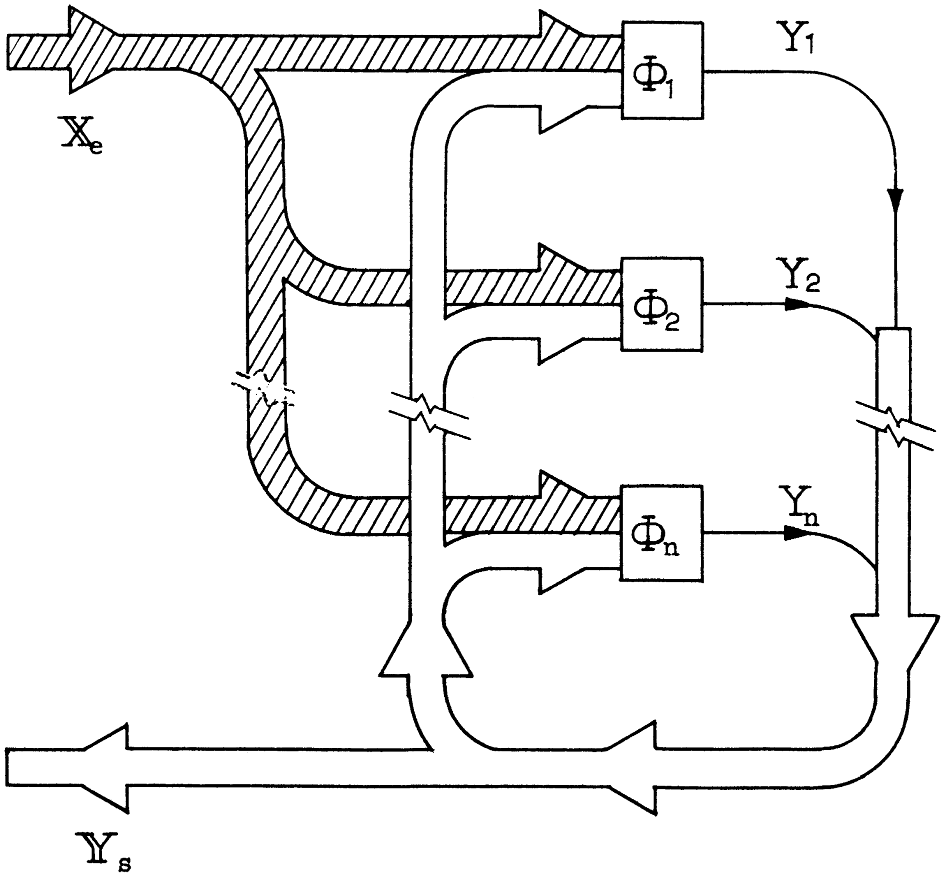
$$\mathbb{Y} = \{ Y_1 , Y_2 , \cdots , Y_{12} \}$$

The only practicable decomposition of the function Φ is one which will yield up each of the Y_i individually. That is, we shall have to reticulate Φ into twelve primitive operators of the form



$i = 1,2,3,\ldots,12$

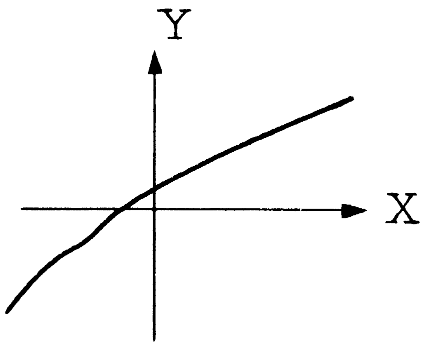
In the most general case, wherein implicit operators are employed, we would thus arrive at the reticulation:



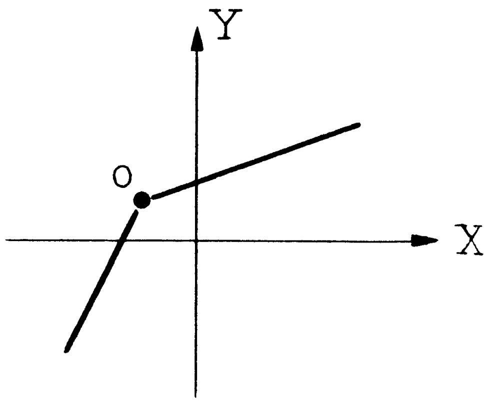
F. Nonlinearity

It is propitious at this point to dwell upon the problem of nonlinearity, and how, in general terms, nonlinear behavioral characteristics are rendered into tractable form for the purposes of quantitative analysis.

We must deal with two types of nonlinearity: (i) curvilinearity, which may be linearized for small excursions about a steady operating point; (ii) essential nonlinearity which cannot be linearized. Examples are sketched below:



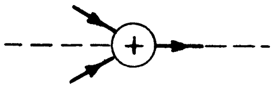
Curvilinear
Characteristic

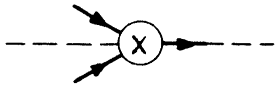


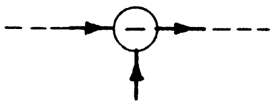
Essentially Nonlinear
Characteristic

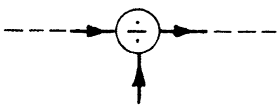
Obviously it is impossible to linearize the essentially nonlinear characteristic in the vicinity of the point O without overlooking a most significant aspect thereof--namely, the discontinuity in slope at O.

Curvilinear characteristics may be approximated to any degree of precision by functions constructed from the basic connectives or operations





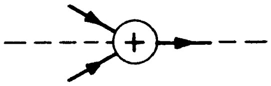


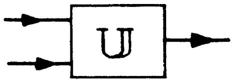


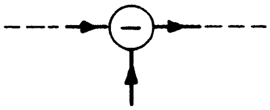
The commutivity of the two inputs in the case of "sum" and "multiply" is most important. However, "minus" and "divide" are asymmetric, and therefor non-commutative. These four operations are the basis for all <u>algebraic functions</u>.

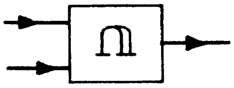
A class of "logical functions" -- the <u>hyperpolyhedral functions</u>-- have been introduced which suffice to construct any linear, curvilinear, or essentially nonlinear characteristic to an arbitrary degree of precision

These are founded upon the operations









Due to the piece-wise linear property of the hyperpolyhedral functions a curvilinear characteristic is automatically linearized, while the discontinuity of an essentially nonlinear characteristic may be exactly preserved.

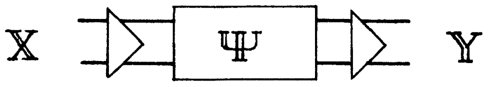
X. Functional Transformations and Computing Functionals

A. Introduction

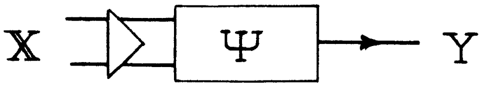
We have said that a multi-ported system which is undergoing some sort of generalized energetic process may be conceived as an element that accepts an input vector $\mathbb{X} = \left\{ X_i \;\middle|\; i = 1, 2, 3, \ldots, n \right\}$ upon which it operates according to the functional $\boldsymbol{\Psi}$ to yield an output vector $\mathbb{Y} \left\{ Y_i \;\middle|\; i = 1, 2, 3, \ldots, n \right\}$; the functional $\boldsymbol{\Psi}$ is such that the entire past state of $\mathbb{X}$ is scanned to yield a single present value of $\mathbb{Y}$. Now, in order to compute the system state at any particular instant it is essential that not only the external outputs $Y_i(t)$ but also the states of each of the internal bonds be computed. We thus desire a reticulation of $\boldsymbol{\Psi}$ which permits each of these internal states, as well as the external output variables to be displayed individually.

B. Computing Functionals

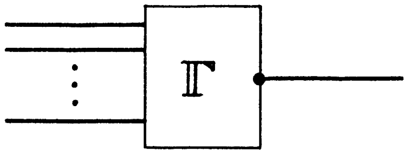
To permit any sort of computing or quantitative description, whether by machine or by hand, the element



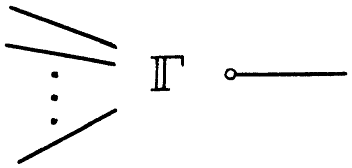
must be reticulated into a set of primitive scalar output functionals



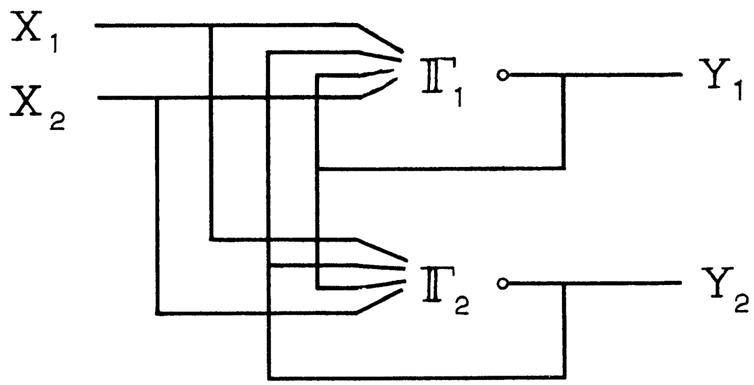
We are thus concerned about the specific way $\boldsymbol{\Psi}$ must be reticulated for computing purposes to allow the state of the system to be completely described. To emphasize that this concern stems from the desire to compute we shall introduce a special symbolism for the primitive computing functional:



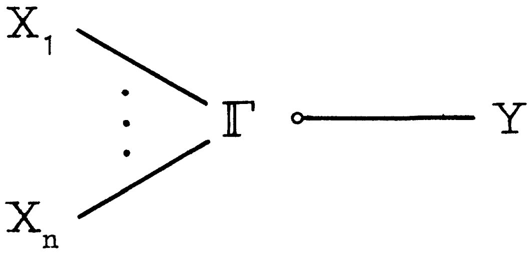
The small circle indicates the output side of the functional. Deleting the box, we have simply



Thus, in order to model a two-ported element two computing functionals might be interconnected as follows:



As is the case in this illustration, it is generally necessary to employ feedback loops in the synthesis of complex computing functionals. As before, the presence of such loops results in <u>implicit</u> computing functionals. Now the functional



represents all possible (i.e., conceivable) deterministic transformations of inputs $\{X_i \mid i = 1, 2, 3, \ldots, n\}$ into an output Y. It determines a present value for Y from the present and all past values of the X_i. Under no circumstances whatever can we demand, nor is there any use for a functional which requires a future value of an input to compute a present Y. Indeed, this may be regarded as the rule or law for the construction of computing functionals.

Consider as two examples of commonly accepted functional operators the ordinary time differentiator and time integrator, symbolized

Derivative: ———— $\mathbb{D}$ ○————

Integral: ———— $\int$ ○————

Now, it is theoretically impossible to compute the "exact" derivative of a variable without a knowledge of its immediate future. That is, by definition,

$$Y(t) = \mathbb{D}[X(t)] = \frac{d}{dt}[X(t)]$$
$$= \lim_{\Delta t \to o} \frac{X(t+\Delta t) - X(t-\Delta t)}{2\Delta t}$$

Thus, according to our conception of an "allowable" computing functional differentiation is not physically realizable.

On the other hand, the operation of running integration is readily constructible for

$$Y(t) = \int [X(t)] = [\int dt] * X(t)$$

requires only a knowledge of the past history of X to compute the present value of Y. Thus, we conclude that, as a rule of thumb, differentiations should be avoided in computing -- indeed there is no way to accurately differentiate -- while a very accurate integrating operation may be physically realized in analog or digital form.

A second important conjugate pair of functional elements are the time advance and time delay:

Advance: ———— $\mathbb{E}$ ○————

Delay: ———— Δ ○————

We immediately note that $\mathbb{E}$ is not an allowable computing functional since

$$Y(t) = \mathbb{E} * [X(t)] = X(t + T)$$

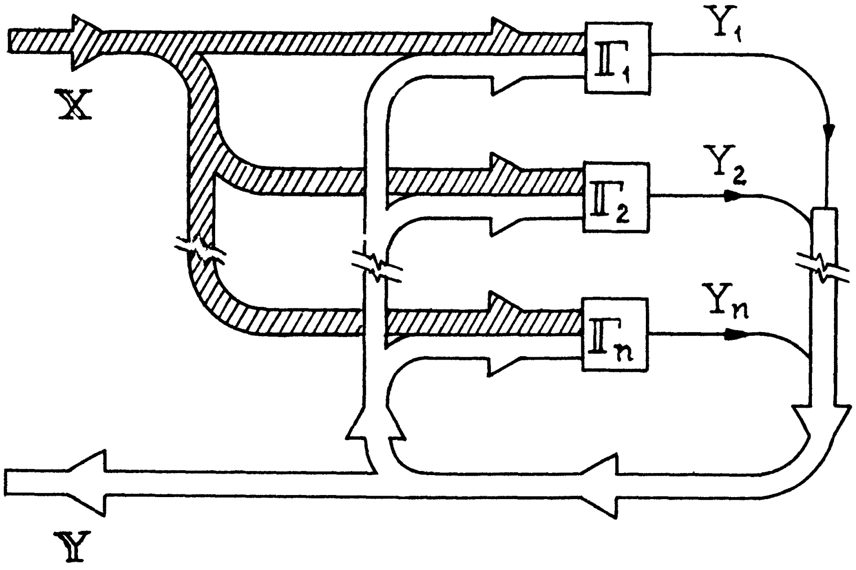
On the other hand, $\mathbb{\Delta}$ is realizable since

$$Y(t) = \mathbb{\Delta} * [X(t)] = X(t - T)$$

Thus, we see, by way of the above examples, that while the relation

$$\mathbb{\Gamma}\mathbb{\Gamma}^{-1} = \mathbb{I} \qquad (\mathbb{I} \sim \text{identity})$$

holds in theory there are many cases where the converse or inverse of a functional is not realizable, and hence the symbol $\mathbb{\Gamma}^{-1}$ has no physical significance. Moreover, although computing functionals are the basic building blocks in any computing program, analog or digital, many functionals may be realized only in one of the two media. For example, precise integration is possible only on the analog machine while precise time delay is a digital operation. However, both the physical model realizable on the analog computer, or the "logical engine" resulting from a digital program may be depicted as shown below:



That is, in general a computing representation or model of a given system may be depicted as a network of computing functionals, containing multiple feedback loops and yielding as outputs all the variables of the system necessary to fix its state.

Background Reading - Computing Systems

(1) HARTREE, D. R. Calculating Instruments and Machines.

This classic work treats the principles underlying both analog and digital machines and describes some of the instruments of historical significance.

(2) IVALL, T. E. Electronic Computers.

A collection of readable British essays on analog and digital electronic devices originally appearing in "Wireless World".

(3) SCOTT, N. R. Analog and Digital Computer Technology.

A contemporary work detailing the structure and applications of modern high speed machines.

(4) von NEUMANN, J. The Computer and the Brain.

A most provocative posthumous essay by the late great mathematician, leaving unanswered the query as to how nature yields such accurate and reliable signals from noisy and erratic components.

XI. Diagrams and the Coding of System Structure

A. Signs

We are here concerned with a problem of communication -- specifically, the transmittal of the description of a reticulated system from one human mind to another. We seek a form of description which is complete yet sufficiently succinct, and of such a nature as to permit a verbal transmittal, over the telephone for example. Thus, an encoding of the schematic description is indicated.

To provide a background for this discussion we consider briefly the general theory of signs or semiotics. Charles Sanders PEIRCE states: "A Sign, or Representamen, is a First which stands in such a genuine triadic relation to a Second, called its Object, as to be capable of determining a Third, called its Interpretant, to assume the same triadic relation to its Object in which it stands itself to the same Object."

All sorts of human communication is accomplished by way of a sign-activity. That is, an individual A employs a sign S to communicate an idea of an object O to a second individual B in whose mind an interpretation I (also a sign) is evolved as a result of perceiving S. The situation is not uncommon in engineering analysis wherein the individuals A and B are the same person, and S is a sketch or diagram drawn as an aid in problem-solving -- a form of self-communication.

Peirce is to be credited with the trichotomy of signs into the classes: (i) Icons; (ii) Indices; (iii) Symbols. Quoting directly from Peirce:

> "A sign is either an icon, an index, or a symbol. An icon is a sign which would possess the character which renders it significant, even though its object had no existence; such as a lead-pencil streak as representing a geometrical line. An index is a sign which would, at once, lose the character which makes it a sign if its object were removed, but would not lose that character if there were no interpretant. Such, for instance, is a piece of mould with a bullet hole in it as a sign of a shot; for without the shot there would have been no hole; but there is a hole there, whether anybody has the sense to attribute it to a shot or not. A symbol is a sign which would lose the character which renders it a sign if there were no interpretant. Such is any utterance of speech which signifies what it does only by virtue of its being understood to have that signification."

Thus, an icon is a characterizing sign which exhibits in and by itself the properties which an object must possess to be denoted by it. Examples of icons are photographs, models, star charts, and chemical diagrams.

An index is a directing sign which refers to its object by a dynamical or spatial connection and otherwise bears no resemblance to the object. Sub- and superscripts, index marks, clocks and meters, and anything which focuses attention or startles may be considered an index.

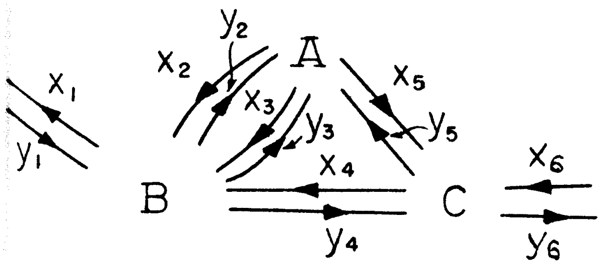
A symbol is a characterizing sign which always involves a rule or convention to establish the connection with the implied object. The utility relies utterly upon the mind of the interpreter to conjure up its meaning and significance. For example, names of people, things, stars, and elements, as well as code marks and mathematical notations, are all symbols.

A sign -- a schematic diagram, for example -- which refers to a physical system as its object, embodies all three classes of sign-action. The bare skeleton of the diagram is iconal, exhibiting directly certain properties of the system. This skeleton, however, is endowed with various labels, arrows, etc. which involve indicial and symbolic sign-action. For example, in a block diagram a component might be labeled "Ψ_1", which directs the reader's attention, or perhaps memory, to the previously made definition of this functional -- as distinguished from the definitions of Ψ_2, Ψ_3, etc. -- and thus involves both indicial and symbolic sign activity.

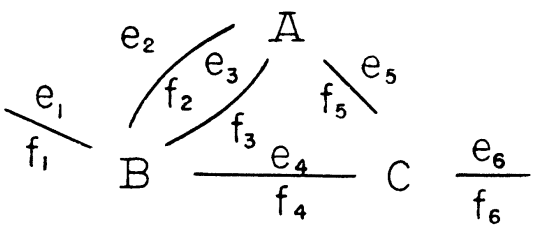
B. Communication of a Computing Structure

Schemata of various sorts -- block diagrams, signal flow graphs, etc. -- are invaluable aids to the description of systems and to the communication of their structure. We are specifically concerned with the problem of describing and communicating the nature of a computing structure, i.e., a network of computing functionals Π_i. We desire a method which is sufficiently flexible to describe the most general types of nonlinear networks and which will lend itself to encoding for the purpose of verbal transmittal.

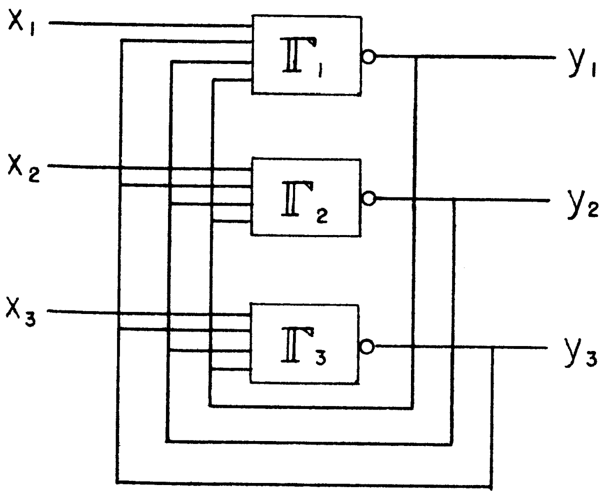
Two essential dichotomies may be discerned in the realm of schematic representations of system structure. The first is now familiar to us: the causal (bilateral signal flow) vs. the non-causal (energy bond) representations. The second dichotomy subdivides the large and variegated class of "branch-node" schemata into, on the one hand, those representations which identify the functional operators with the nodes and the signal variables with the branches (block diagrams); and, on the other hand, those representations which identify the variables with the nodes and the operators with the branches (Mason-Tustin signal flow graph).



Causal Bilateral Signal Flow Diagram



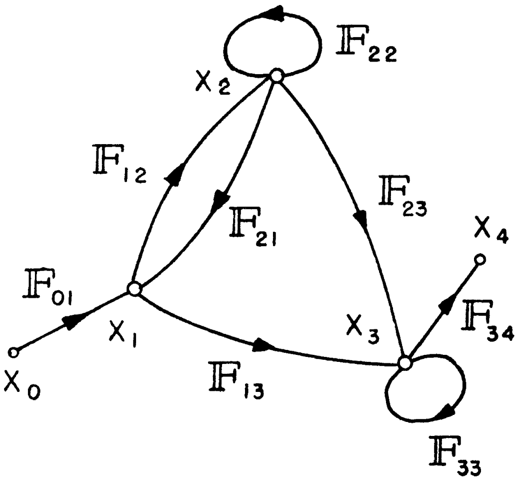
Non-Causal Energy Bond Diagram



Functional Block Diagram

Operators: Nodes

Variables: Branches



Signal Flow Graph

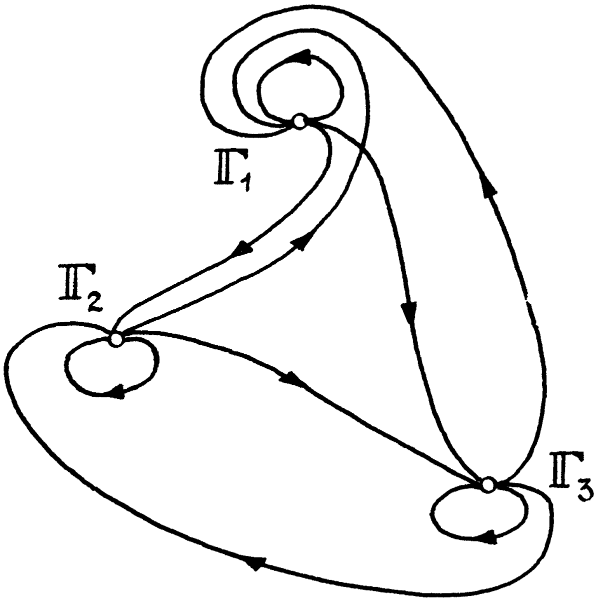
Variables: Nodes

Operators: Branches

The non-causal representation, a generalized circuit diagram, uncluttered and simple, enables the experienced analyst to visualize quickly the behavior of a system, while the causal description is essential for a detailed quantitative understanding of its performance. The block diagram is especially suited to determining the transfer characteristic of a structure of interconnected elements, provided the boundaries of the elements have been correctly chosen. In the case of a computing structure, which is our present concern, these boundaries are generally self evident. The block diagram has the distinct advantage of being applicable to the case of nonlinear as well as linear systems. The signal flow graph, on the other hand, may be used precisely only to describe linear networks since a summary action is implied at each of the nodes; that is, for example

$$x_1 = \mathbb{F}_{o1}\, x_o + \mathbb{F}_{21}\, x_2$$

For all these cases, however, we seek a representation which is capable of being encoded, and for this purpose the following branch-node structure suggests itself:



But this structure may be easily encoded by way of the following tabulation:

Y	Γ	X
1	Γ_1	1,2,3
2	Γ_2	2,3,1
3	Γ_3	3,1,2

Corresponding to each node there is a single output y, that results from the operation of the associated functional Γ upon the input signals, which in this case are simply the outputs of all three nodes. Thus, for example, the first row of the table might be read, "the signal y_1 results from the operation of Γ_1 upon y_1, y_2, and y_3". In actuality, of course, the entries in the Γ-column would indicate the nature of the functionals, say by way of a numerical coding: 1 for an upper selector, 2 for a lower selector, 3 for an integrator, etc. It is thus possible to communicate succinctly a complex structure in the form of a table or sequence of numbers only. The task of transforming this number sequence into a readable diagram and vice versa is almost trivial.

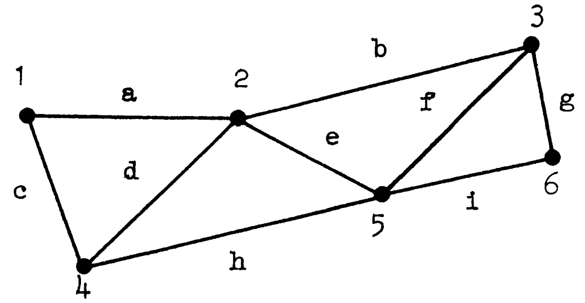
What we have done here is to treat a specific application of the broader *theory of graphs*, which in turn stems from the mathematical discipline of *combinatorial topology*. This general study deals with the ways in which the structural connexity of a space may be described and communicated; we recognize this as precisely the problem with which we have been concerned, wherein "the space" happens to include a computing system and the connectedness of interest to us embraces the functional relationships between the several computing components. In combinatorial topology connexity is communicated by way of *incidence matrices*, a condensed form of which are the coded tables here suggested for use in communicating system structure.

C. Combinatorial Topology - Incidence Matrix

A. W. TUCKER states: "Topology deals with the rudimentary geometrical properties which depend on continuity rather than on size and shape." The domain of discourse is a *space* in which the topologist attempts to establish theorems related to connexity and structure.

Henri POINCARÉ is generally cited as the originator of this branch of mathematics, which he named analysis situs.

Connexity is depicted by way of linear graphs or, alternatively, by incidence matrices. A linear graph is constituted from nodes and branches. A digraph (directed graph) is a linear graph in which the branches have been endowed with a directional sense. An example of an ordinary linear graph is given below:



In this graph there are nine branches and six nodes. The associated incidence matrix may be easily written:

	a	b	c	d	e	f	g	h	i
1	1	0	1	0	0	0	0	0	0
2	1	1	0	1	1	0	0	0	0
3	0	1	0	0	0	1	1	0	0
4	0	0	1	1	0	0	0	1	0
5	0	0	0	0	1	1	0	1	1
6	0	0	0	0	0	0	1	0	1

In this matrix an entry of "1" indicates a branch-node impingement, while an entry of "0" indicates no impingement. The elements are therefore labelled incidence numbers.

A topological space is a complex constituted from a number of simplexes or cells; these are labelled, according to convention, as follows:

0-cells : nodes
1-cells : branches
2-cells : loops

Hence, the incidence matrix discussed above, which depicted a node-branch structure, is called the "01" incidence matrix, or simply $\mathbb{I}_{01}$. Poincaré defined the numbers

$$a_k = \text{number of k-cells in a complex}$$

$$a_o = \text{number of 0-cells}$$
$$a_1 = \text{number of 1-cells}$$
$$a_2 = \text{number of 2-cells}$$

The rank of the incidence matrix $\mathbb{I}_{k,k+1}$ is denoted r_k. Since no significance has been attributed to $\mathbb{I}_{k,k+1}$ for $k=-1$ it is necessary to restrict this definition to hold only for $k = 0,1,2,\ldots$. Hence, we say that

$$r_k = \text{rank of } \mathbb{I}_{k-1,k} \text{ (for } k = 0,1,2,\ldots) \; ; \; r_{-1} \equiv 0$$

We also define the k^{th} order Betti number

$$b_k = a_k - r_k - r_{k-1}$$

so that, in particular, the zeroeth and first Betti numbers are given by

$$b_o = a_o - r_o \quad ; \quad b_1 = a_1 - r_1 - r_o$$

which requires that some significance be attached to b_o. Accordingly, we define

$$b_o \equiv \text{number of separate connected parts in a complex.}$$

With this it is now convenient to define the rank R of the linear graph as

$$R \equiv r_o = a_o - b_o$$

which yields an alternative definition of the first Betti number for linear graphs, since $r_1 \equiv 0$, namely

$$b_1 = a_1 - a_o + b_o = a_1 - R$$

It is also propitious to observe that some authors refer to the first Betti number as the nullity, N.

The Euler characteristic is defined in terms of either the b_k or the a_k as follows:

$$K \equiv \sum_k b_k \quad (-1)^k \quad \equiv \sum_k a_k \quad (-1)^k$$

The Euler characteristic for a connected linear graph of V nodes and B branches is simply

$$K = b_o - b_1 = a_o - a_1$$

or

$$K = 1 - N = V - B$$

Since $R = V - 1$ we thus obtain the fundamental invariant relation for all linear graphs

$$\boxed{B = R + N}$$

which is identical to the previous result $b_1 = a_1 - R$.

By way of illustration of the significance of some of the above characteristic numbers three theorems are stated.

Theorem 1. If we start with the 0-cells of a linear graph and add the 1-cells one by one, the number of 1-cells added joining nodes not previously connected is r_o and the number of 1-cells added joining vertices already connected is b_1.

In connection with this theorem it is well to point out that a complex which contains no loops -- i.e., no closed paths within the structure but which would contain a loop with the addition of a single branch, is called a tree. A forest is a complex consisting of a number of disconnected trees.

Theorem 2. The first Betti number of a forest is zero.

Theorem 3. If the first Betti number of a graph is b_1, we can remove b_1 1-cells from it, but no fewer, which will reduce it to a forest.

These theorems are stated without proof for the purpose of illustration only. From them we observe the importance of the rank R and nullity N in the topological characterization of a space.

Background Reading - Signs

(1) PEIRCE, C. S. Philosophical Writings, (edited by J. Buchler), Logic as Semiotic: Theory of Signs.

Peirce presents his form of the theory of signs--the logic of semiotic. Much of the point of view adopted in this course originates with Peirce, although this subject has been taken up and colored by more recent thinkers in this field (and occasionally presented in more readable fashion).

(2) GALLIE, W. B. Peirce and Pragmatism

Gallie presents a compact summary of Peirce's semiosis and theory of signs.

(3) YOUNG, J. W. Lectures on Fundamental Concepts of Algebra and Geometry, pp. 226-239, (Growth of Algebraic Symbolism, by U. G. Mitchell)

The history of the use of symbols in algebra and arithmetic is traced.

(4) MORRIS, C. W. Foundations of the Theory of Signs

Morris presents (without adequate citation) much of Peirce's thought on this subject.

(5) CHERRY, C. On Human Communication, Chap. 3, pp. 219-226.

This is a modern text in which signs are discussed as a part of the broader subject of communication. Much of Peirce's thought is again represented.

(6) TRUXAL, J. G. Automatic Feedback Control System Synthesis, Chap. 2.

A discussion is given of the disadvantages of block diagrams and the Mason signal flow graph is presented as a useful tool in systems analysis.

Background Reading - Topology

(1) SYNGE, J. L. The Fundamental Theorem of Electrical Networks, Quarterly of Applied Math., July 1951, p. 113.

In his development of the theorems and concepts leading up to the "fundamental theorem" the author employs a very readable intuitive approach. Much of this development is purely a discussion of topology and digraphs which is direct support of the material on this subject presented herein.

(2) TUCKER, A. W. The Topological Concept of Space. (A lecture given at the Galois Institute of Mathematics).

Tucker discusses many of the essential concepts of topology without resorting to formal mathematical proofs. Thus, his approach lends itself to a deepening insight into this subject, beyond the superficial statements made in these notes.

(3) SINGER, James. One-Dimensional Analysis Situs. (A lecture given at the Galois Institute of Mathematics). (1935)

This reference contains much of the material used in these notes. The theorems merely stated herein are stated and proved by Singer, as are several additional theorems which concern the structure and connexity of linear graphs.

(4) SINGER, James. Two-Dimensional Analysis Situs. (A lecture given at the Galois Institute of Mathematics). (1936)

Many of the statements made by Tucker are discussed more thoroughly in this reference which extends, along intuitive lines, into the topology of two-dimensional spaces (surfaces).

D. Coded Representation of Graphs and Digraphs

The original branch-node incidence matrix of the previous section may be encoded in a simple array merely by condensing or collapsing either rows or columns in the following alternative fashions:

ROW CODE	
1	a c
2	a b d e
3	b f g
4	c d h
5	e f h i
6	g i

COLUMN CODE	
a	1 2
b	2 3
c	1 4
d	2 4
e	2 5
f	3 5
g	3 6
h	4 5
i	5 6

It is readily apparent that an encoding by rows gains rapidly in efficiency and simplicity as the connexity of the structure increases if the specific node and branch tags are both to be transmitted. Nevertheless. we shall have frequent occasion to use both forms of coding as required.

Second (Branch-Loop) Incidence Matrix

In addition to the first (node-branch) incidence matrix, the matrix indicating the cyclic or closed-loop character of the system structure is also of fundamental topological interest. This circuital or branch-loop incidence may be determined for any reticulate system by indicating the incidence of all branches upon N + 1 independent loops where N is the nullity (i.e. the number of branches-out-of-tree) of the structure.

		LOOPS				
		I	II	III	IV	V
BRANCHES	a	1	1	.	.	.
	b	1	.	.	1	.
	c	1	1	.	.	.
	d	.	1	1	.	.
	e	.	.	1	1	.
	f	.	.	.	1	1
	g	1	.	.	.	1
	h	1	.	1	.	.
	i	1	.	.	.	1

Dual Graphs

For the graph previously depicted and discussed, the rank $R = 5$ and the nullity $N = 4$. Therefore, for the dual graph, the rank $R^* = 4$ and the nullity $N^* = 5$. This dual graph may be constructed directly from the transpose of the second incidence matrix, merely using the topological dual isomorphism:

$$(N + 1) \text{ Loops} \longleftrightarrow (R^* + 1) \text{ Dual Nodes}$$

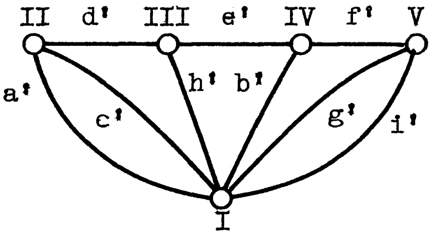
$$(B) \text{ Branches} \longleftrightarrow (B^*) \text{ Dual Branches}$$

Thus the first incidence matrix of the dual graph is merely the transpose of the second incidence matrix of the original graph. This gives, in coded form:

I a' b' c' g' h' i'
II a' c' d'
III d' e' h'
IV b' e' f'
V f' g' i'

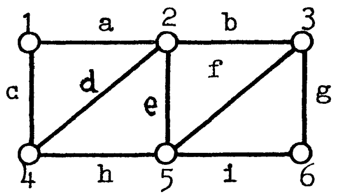
corresponding to the graphical form:

DUAL:
$R^* = 4$
$N^* = 5$
$B^* = 9$



by contrast to the original figure:

ORIGINAL:
$R = 5$
$N = 4$
$B = 9$



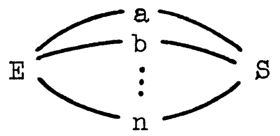
A well-known theorem of topology due to Hassler WHITNEY states that a dual graph can exist only for a planar graph (i.e. a linear graph which can be homeomorphically mapped onto a plane or a sphere).

E. Coding the Energetic Structure of Multiport Systems

The previous incidence matrices and equivalent codes may be used for the topological structuring of multiport systems, provided that the system is closed, and the following correspondence is employed:

MULTIPORT ELEMENTS ⟷ NODES
POWER BONDS ⟷ BRANCHES

First, it is possible to close all otherwise open multiport systems by a simple artifice. Since an n-ported system S must necessarily be bonded to an n-ported environment E, we can always annex the environment to the system itself to form a necessarily closed system, in the fashion:



To emphasize the complementary aspect of the environment in this circumstance we may denote the environment of S by the underscored symbol, $\underline{S}$. Thus the closed system becomes

The duality between S and $\underline{S}$ is complete since

$$\underline{\underline{S}} \equiv S$$

which means in words that the environment of the environment of a system is the system itself.

Thus if the system S is coded as:

S a b c ... n

then we may designate the environment $\underline{S}$ of S as

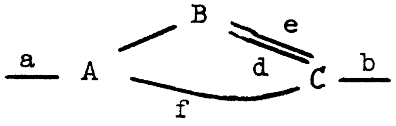
$\underline{S}$ a b c ... n

and the two subsystems as a single closed system becomes in coded form:

S a b c ... n
$\underline{S}$ a b c ... n

If the system is closed to begin with no complementary element, $\underline{S}$, is required for closure.

In the nontrivial case the internal energetic reticulation is also given. For example the structure:



might be encoded as

$\underline{S}$ a b
A a c f
B c d e
C b d e f

From the code itself we may infer the following facts, among others:

1) The overall system is a 2-port
2) It has been reticulated into 3 multiports, namely
    a) two 3-ports
    b) one 4-port

The assignment of CAUSALITY may be accomplished using the code alone as follows:

I ... EFFORT inputs are unmarked;
II ... FLOW inputs are underscored.

Thus a typical causality would be as follows:

$\underline{S}$ $\underline{a}$ b
A a $\underline{c}$ f
B c $\underline{d}$ $\underline{e}$
C $\underline{b}$ d e $\underline{f}$

This means that S itself is of the form

S a $\underline{b}$

since $\underline{S\ a\ \underline{b}} \equiv \underline{S}\ \underline{a}\ b$ is its complement.

The Mechanics of System Interconnection

Consider that we were assembling the original system from the subsystems A, B, C, whose ports have been assigned consistent causality. In this case we would have started with

$$A \quad a \ \underline{b} \ c \quad ; \quad B \quad a' \ \underline{b'} \ \underline{c'} \quad ; \quad C \quad \underline{a''} \ c'' \ \underline{d''}$$

The primes would not generally appear in the separate listings and are here indicated only to prevent confusion.

The initial step requires the unique labelling or, better, numbering of all ordered ports, for example as follows:

$$A \ 1.\underline{2}.3. \qquad B \ 4.\underline{5}.\underline{6}. \qquad C \ \underline{7}.8.9.\underline{10}$$

This array corresponds to the (element-bond) incidence matrix

$$\begin{array}{c|cccccccccc|}
 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
\hline
A & 1 & \underline{1} & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
B & \cdot & \cdot & \cdot & 1 & \underline{1} & \underline{1} & \cdot & \cdot & \cdot & \cdot \\
C & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \underline{1} & 1 & 1 & \underline{1} \\
\hline
\end{array}$$

The particular interconnections given previously may now be expressed by

$$\begin{array}{ccc}
2 & = & 4 \\
5 & = & 8 \\
6 & = & 9 \\
3 & = & 10
\end{array}$$

These column identifications result in column additions in the system matrix and reduce the matrix to:

$$\begin{array}{c|cccccc|}
 & & 4 & 10 & 8 & 9 & \\
 & & + & + & + & + & \\
 & 1 & 2 & 3 & 5 & 6 & 7 \\
\hline
A & 1 & \underline{1} & 1 & \cdot & \cdot & \cdot \\
B & \cdot & 1 & \cdot & \underline{1} & \underline{1} & \cdot \\
C & \cdot & \cdot & \underline{1} & 1 & 1 & \underline{1} \\
\hline
\underline{S} & \underline{1} & \cdot & \cdot & \cdot & \cdot & 1 \\
\hline
 & a & b & c & d & e & f
\end{array}$$

where the environmental $\underline{S}$ row may be added such that the column sum vanishes identically for every column.

The corresponding coded system may now be written:

$\underline{S}\ \underline{a}\ f$

$A\ a\ \underline{b}\ c$

$B\ b\ \underline{d}\ \underline{e}$

$C\ \underline{c}\ d\ e\ \underline{f}$

To render this in a coded form *identical* to that of the original, only simple permutation of letters is required in the form:

$$\downarrow\ \begin{array}{ccc} f & b & c \\ \hline b & c & f \end{array}$$

This yields the equivalent code

$\underline{S}\ \underline{a}\ b$

$A\ a\ \underline{c}\ f$

$B\ c\ \underline{d}\ \underline{e}$

$C\ \underline{f}\ d\ e\ \underline{b}$

which is merely a permutation of the first system and is therefore topologically or structurally identical.

Background Reading -- Graphs, Digraphs, and Networks

(1) CAYLEY, A. On the Analytical Forms called Trees, with Application to the Theory of Chemical Combinations, *Report of the British Association for the Advancement of Science*, pp. 257-305 (1875)

(2) KEMPE, A. B. A Memoir on the Theory of Mathematical Form, *Philosophical Transactions*, pp. 1-70 (1886).

A little known and truly remarkable anticipation of combinatorial topology whose origin is usually credited to the papers of POINCARE.

(3) KOENIG, D. *Theorie der Endlichen und Unendlichen Graphen*, Chelsea Publishing Co., New York (1950).

This relatively recent book has now become a classic in this field.

Background Reading -- Graphs, Digraphs, and Networks (continued)

(4) WHITNEY, H. Non-separable and Planar Graphs, Transactions of the American Mathematical Society, Vol. 34, pp. 339-362 (1932).

The author here proves for the first time that duals exist only for planar graphs and therefore for planar logical and electrical networks.

(5) HOHN, F. E., S. SESHU, and D. D. AUFENKAMP. The Theory of Nets, Transactions of the IRE, Vol. EC-6, No. 3, pp. 154-161 (September, 1957).

The authors generalize the concept of a digraph into a net to include certain higher order structural information. Many theorems and properties of universal value may then be adduced.

(6) SHIMBEL, A. Structure in Communication Nets, Proceedings of the Symposium on Information Networks, Polytechnic Institute of Brooklyn, Brooklyn, pp. 199-203 (1954).

This paper propounds concepts and methods which enable the determination of the minimum paths and resultant trees in any communication digraph.

(7) HARARY, F. Structural Duality, Behavioral Science, Vol. 2, No. 4, pp. 255-265 (October, 1957).

A very readable treatment of various duality transformations applied to graphs and digraphs.

(8) GRAYBEAL, T. D. Block Diagram Network Transformation, Electrical Engineering, pp. 985-990 (November, 1951).

(9) STOUT, T. M. A Block-Diagram Approach to Network Analysis, AIEE Transactions, pp. 255-260 (November, 1952).

(10) MASON, S. J., Feedback Theory--Some Properties of Signal Flow Graphs, Proc. Inst. Radio Engrs. 41, 1144-1156 (September, 1953).

(11) .... Feedback Theory--Further Properties of Signal Flow Graphs, Proc. Inst. Radio Engrs. 44, 920-926 (July, 1956).

The four papers above deal with transformations and equivalencies of flow graphs.

(12) SESHU, S. and REED, M. B. Linear Graphs and Electrical Networks .

This excellent text has a summary of much of the above and an excellent bibliography.

XII. State-Determined Systems

A. Introduction

Generalized Dynamics and State-Determined Systems

A class of one-port elements of great practical importance is associated with the historical concept of state-determined systems. For such idealized systems the specification of a certain delimited set of parameters is sufficient to describe completely the behavior of the system. These parameters are said to determine the state of the system in the sense that when the values of such parameters are known at any instant, the behavioral configuration is also known at that instant.

Any particular system is then specified merely by fixing a set of static, generally nonlinear relationships between these state variables. All problems in the generalized dynamics of such state-determined systems are reduced to the purely kinetic or kinematic "motion" of a representative point in a multidimensional abstract space, called the phase space. In such a phase space any single point represents a possible state of the system, and a connected set of such points, or phase trajectory, represents a "history" of the system.

The essential significance of the concept of a state-determined system rests in the fact that the future behavior of such a system is determined completely in terms of the complete specification of the instantaneous present state of the system, together with the temporal fluctuations of all "external forces" during the future period. In many practically important cases where the external effects are small enough to be neglected, so that the system may be regarded as effectively isolated or closed, the future behavior is determined once and for all upon specification of the initial state alone.

This fascinating notion utterly dominated the growth and evolution of classical mechanics until just before the outset of the present century. Indeed, we may quote the following from G. D. BIRKHOFF (Dynamical Systems--1927):

> In dynamics we deal with physical systems whose state at time t is completely specified by the values of n real variables
>
> $$x_1, x_2, x_3, \ldots, x_n$$

> Accordingly, the system is such that the rates of change of these variables, namely .
>
> $$dx_1/dt, \ dx_2/dt, \ dx_3/dt, \ \ldots, \ dx_n/dt$$
>
> merely depend upon the values of the variables themselves, so that the laws of motion can be expressed by means of n differential equations of the first order,
>
> $$dx_i/dt = X_i\,(x_1, x_2, x_3, \ldots, x_n) \quad (i = 1, 2, 3, \ldots, n)$$

However, more recent physical treatments have acknowledged the fact that no material system is ever isolated, nor is the instantaneous state ever capable of complete determination; out of this realization have been evolved modern statistical mechanics and other stochastic views of the "knowable" physical world. Despite this, state-determined "models" of material systems will continue to play an extremely useful role in engineering analyses which are directed toward the practical prediction of approximate performance.

State-Determined Systems as Models

Quoting from the paper by Arturo Rosenblueth and Norbert Wiener, The Role of Models in Science:

> No substantial part of the universe is so simple that it can be grasped and controlled without abstraction. Abstraction consists in replacing the part of the universe under consideration by a model of a similar but simpler structure. Models, formal or intellectual on the one hand, or material on the other, are thus a central necessity of scientific procedure.

Thus, whenever an engineering problem must be studied, other than by direct manipulation or experimentation with the actual system involved, it is necessary to have recourse to models of some type.

Often these are real or actual models, in which physical counterparts are involved. Sometimes they are earlier versions of the same type of system or are extant and similar devices. Often they are simplified or scaled down versions in the form of research models and pilot plants. Again, physical models in the form of analogies are frequently used with considerable effectiveness.

In many other circumstances, only conceptual models are employed, in which rational abstractions and idealizations are made to correspond, with more or less validity, to real situations. These may be rendered precise--but not necessarily accurate--in the form of <u>mathematical models</u>, in which the component elements are mathematical variables interrelated through various mathematical operations.

If these operations can be restricted to a small collection of simple computing operations, we can thus construct a <u>computing model</u> corresponding, at least approximately, to any given engineering situation. Lastly, if the operations upon and interconnections between variables are actually <u>realized</u> in a physical device, the corresponding computing system does, in fact, constitute a physical model

In previous chapters we have discussed the nature and description of computing models in considerable detail. We now turn our attention to the <u>state-determined</u> model of the physical universe, which is intrinsically a mathematical model, and, although it has certain shortcomings which will be pointed out, it does succeed in describing a great variety of phenomena.

B. Elements of a State-Determined System

Any state-determined system may be reticulated into just two kinds of multi-ported elements: (i) one-port impedances, which are generically denoted -X, and which include ideal resistances (-R), capacitances (-C), and inertances (-I) together with the ideal effort source, (-E), and ideal flow source, (-F); (ii) multi-ported energy junctions, which are generically denoted $\cdot\underset{\cdot}{J}\cdot$, and which include the flow junction $\cdot\underset{\cdot}{0}\cdot$ and the effort junction $\cdot\underset{\cdot}{1}\cdot$. Hence, we have at our disposal a field of seven elements:

1 Port Impedances -X	3 Port Energy Junctions $\cdot\underset{\cdot}{J}\cdot$
-R	
-C	$\cdot\underset{\cdot}{0}\cdot$
-I	
-E	$\cdot\underset{\cdot}{1}\cdot$
-F	

The dynamical models of Newton and Faraday and the field model of Maxwell for electro-mechanical interaction may all be reconstructed from this seven-element universe. Moreover, Lagrangian and Hamiltonian mechanics treats systems which may also be reticulated into these same elements. Hence, at first glance, we might be prone to say that all the universe may be modeled as a network of these seven multi-ports!

A further reflection will reveal, however, that there are certain two-and three-port elements which for example satisfy the condition $\sum \mathbb{P} \equiv 0$ yet are not energy junctions, namely, ideal transducers, transformers, levers, and differentials. Obviously, without such elements it is impossible to model or to construct a host of essential man-made devices. Thus, we must conclude that the seven-element universe of classical mechanics is by no means complete. To represent all conceivable systems we must not only add other state-determined elements, but also augment these with more general multi-port elements.

C. The Mathematical Construct of a State-Determined System

We shall now concern ourselves with the mathematical structure of the state-determined model. The most general form consistent with the underlying assumptions and concepts of state-determinism will be compared with the classical form which was stated briefly by way of the quotation above.

First of all, the state-determined system is one whose condition at any time (t) is precisely and completely specified by a finite set of variables $\mathbf{X}(t) = \left\{ x_i(t) \mid i = 1,2,\ldots,n \right\}$. It is therefore easy to see that an instantaneous condition or state of the system is represented by a point in an n-dimensional phase-space. As the condition of the system alters, due to the action of disturbances at its boundaries the state-vector $\mathbf{X}$ traces a path in the phase space which we appropriately call the phase trajectory.

Before proceeding further it is well to point out that a system whose condition is completely described only by specifying an <u>infinite</u> number of variables cannot be "state-<u>determined</u>"-- indeed there is no way of computing its state in a finite time.

We have said that the state-vector $\mathbf{X}$ is a function of time, t;

more generally, it varies according to some running parameter which usually possesses the dimension of time. Actually there is no "clock" capable of keeping a perfectly continuous running record of time; however, we may think of the parameter t as being either discrete or continuous for the purpose of computation. In particular, analog computation usually employs a continuous time parameter (either "real", or scaled), while digital computation uses a discrete time parameter, i.e., the state-vector is computed every hundredth of a second, for example.

With these notions understood we may now turn to the primary assumption in the construction of the state-determined model. This concerns the manner in which the instantaneous state is related to earlier states. This result is equivalent to restricting the nature of the phase trajectory which may pass through a certain state, say $\mathbb{X}(o)$. In the most general form, this condition is expressed as follows, in accordance with the functional notation we have employed up to this point:

$$\mathbb{X} = \mathbb{T} * \Phi * \mathbb{X}$$

That is, in words, a unique present state $\mathbb{X}(t)$ is determined by the trajectory $\mathbb{X}(t-\tau | \tau > o)$. Such a determination involves a static (possibly implicit) function Φ of all the variables $\{x_i \mid i = 1,2,\ldots,n\}$ and a time translation operator $\mathbb{T}$ of a peculiar type. Let us immediately compare this purportedly general form with the classical form which we have already seen, namely wherein the rate of change of the state-vector $\mathbb{X}$ is a static function of the state variables $\{x_i \mid i = 1,2,\ldots,n\}$. That is,

$$\frac{d\mathbb{X}}{dt} = \Phi \{(x_i \mid i = 1,2,\ldots,n)\}$$

This we may write in the equivalent compact form

$$\frac{d\mathbb{X}}{dt} = \Phi(\mathbb{X})$$

or, in the integral form

$$\mathbb{X} = [\int_{-\infty}^{t} dt] * \Phi(\mathbb{X}) = \mathbb{T} * \Phi * \mathbb{X}$$

Upon comparison with the form first stated it is evident that the time translation operator $\mathbb{T}$ corresponds to, and therefore must be of the same general nature as the running integrator [$\int dt$].

D. The Variables of State in Generalized Dynamics

Newton founded his axiomatic dynamics upon the concepts of force and momentum; thus, a system of mass points was viewed as connected by way of force interactions, and the motion of each particle was determined according to the fundamental law:

[Vector summation of forces = time rate of change of momentum.]

The forces active on a given particle could arise either from disturbances of an origin external to the system, or from internal interactions.

Newton's second law is often loosely stated in the form

$$\sum F = ma$$

rather than the original and correct form

$$\sum F = \frac{d}{dt}(mv)$$

Following Newton's enuciation of his "laws of motion" there ensued a battle royal among the learned dynamicists of the next two centuries over the notion of "force" and its true usefulness in the dynamical description of a system. Hertz, in his introduction to The Principles of Mechanics, summarizes these arguments in a most profound manner, and discusses the various "images" of a generalized dynamical situation.

In our position, however, it is perfectly acceptable to define a certain quantity, p , as the generalized momentum of an element, which is related to the associated generalized force or effort, e , according to the exact law

$$e = \frac{dp}{dt} \quad \text{or} \quad p = \int e dt$$

This form is in agreement with Newton's original statement, and is acceptable also in the light of relativistic mechanics which tell us that there is, in reality, a functional dependency between momentum and velocity such that as the speed of light is approached the momentum becomes infinite.

There arose, as a result of the work of Lagrange and Hamilton, another image of a dynamical situation wherein it was the potential and kinetic energies which defined the state of the system. We must realize in connection with this that there is always a correspondence between the total strain or displacement of a system and its potential energy, which may arise as a result of internal deformations as well as gross displacements in a force or potential field. Maxwell points out that the correspondence is usually not a simple one, particularly in the case of the energy of deformation. On the other hand, there is a simple relationship which connects the kinetic energy of a particle and its velocity. Thus, it might be said that the two state variables which arise out of this energetic description are the generalized displacement, q , and the generalized velocity or flow, f . These are, of course, connected by the relationship

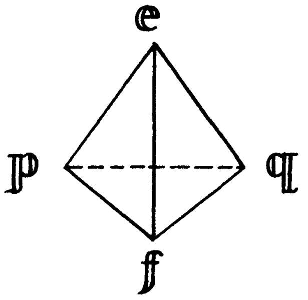
$$f = \frac{dq}{dt} \qquad \text{or} \qquad q = \int f dt$$

The purpose of the above discussion is simply to justify the selection of the four variables e , p , f , and q as the variables of state with which the dynamical or energetic condition of any physical state-determined system may be described. Thus, for a multi-degree of freedom system the following state vectors would be employed:

$$\mathbf{e} = (e_i \mid i = 1,2,3,\ldots,n) \qquad \mathbf{f} = (f_i \mid i = 1,2,3, \ldots,n)$$
$$\mathbf{p} = (p_i \mid i = 1,2,3,\ldots,n) \qquad \mathbf{q} = (q_i \mid i = 1,2,3,\ldots,n)$$

E. The Tetrahedron of State

It is now possible to identify each of the four state variables with a vertex of a "tetrahedron of state" and consider a given system as characterized by the functional relationships between the variables, these being associated with the edges of a tetrahedron:



The correspondences between $\mathbb{e}$ and $\mathbb{p}$ and between $\mathbb{f}$ and $\mathbb{q}$ are, of course, implicit in the construct itself; that is,

$$\mathbb{p} = \int \mathbb{e}\, dt \quad \text{or} \quad \mathbb{e} = \frac{d\,\mathbb{p}}{dt}$$

$$\mathbb{q} = \int \mathbb{f}\, dt \quad \text{or} \quad \mathbb{f} = \frac{d\mathbb{q}}{dt}$$

However, the other essential correspondences, which are but three in number, are static vector functions peculiar to a given system, namely,

$$\mathbb{e} = \mathbf{\Phi}_R(\mathbb{f}) \quad \text{or} \quad \mathbb{f} = \mathbf{\Phi}_G(\mathbb{e})$$

$$\mathbb{q} = \mathbf{\Phi}_C(\mathbb{e}) \quad \text{or} \quad \mathbb{e} = \mathbf{\Phi}_S(\mathbb{q})$$

$$\mathbb{p} = \mathbf{\Phi}_I(\mathbb{f}) \quad \text{or} \quad \mathbb{f} = \mathbf{\Phi}_\Gamma(\mathbb{p})$$

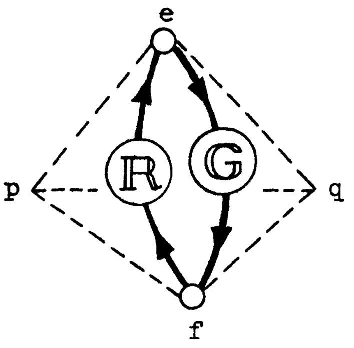
Thus, we see that the characterization of all dynamical or time-dependent interactions is embodied in the $\mathbb{p}$ - $\mathbb{e}$ and $\mathbb{q}$ - $\mathbb{f}$ relations, while the remaining relationships are of a simple static nature. These latter will now be dealt with in detail.

F. The Characteristic Static Relations

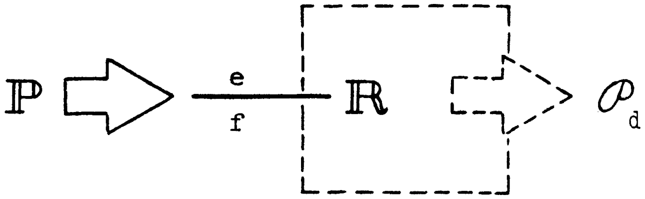
Before discussing in any detail the static relationships introduced in the previous section, it is well to emphasize that the structure of these relations is peculiar to a given system; it is, in fact, derived from the reticulation of that system into one-port impedance elements and multi-ported energy junctions. Thus, for example, the form of the static vector function $\mathbf{\Phi}_r$, or its dual, $\mathbf{\Phi}_g$ is determined by the disposition of resistive elements (-R) in the reticulated system. Similarly, $\mathbf{\Phi}_c$ and $\mathbf{\Phi}_i$ are determined according to the disposition of capacitive and inertive elements, respectively.

Below we emphasize the expected (indeed inevitable!) nonlinearity of these relations. Also, in each case the dualism of relationships is recognized, along with the resultant necessity to distinguish between two complementary "energies" associated with each impedance element. Failure to make this careful distinction in the presence of nonlinear relationships between the state variables has, at times, resulted in serious and substantial errors in the calculation of energy storage and dissipation terms.

1. Resistance-Conductance Relations and Generalized Energy Dissipation



State Relations



Pure Energy Dissipation

$$\boxed{\mathbb{P}(t) \equiv \mathcal{P}_d}$$

For any one-port element the generally nonlinear static relationship between effort (e) and flow (f) can always be considered as a:

RESISTANCE
Relationship
$\mathbb{R}$

or

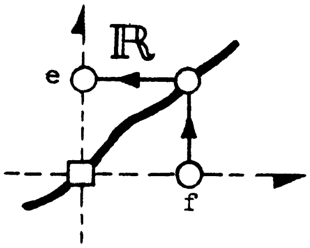
CONDUCTANCE
Relationship
$\mathbb{G}$

We will distinguish between these two converse modes, particularly when we consider causal sense, as follows:

$f \rightarrow \mathbb{R} \rightarrow e$

GENERALIZED
RESISTANCE

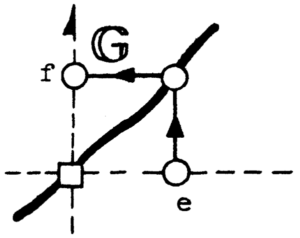
$$e = \Phi_{ef}(f) = \mathbb{R}(f)$$



$e \rightarrow \mathbb{G} \rightarrow f$

GENERALIZED
CONDUCTANCE

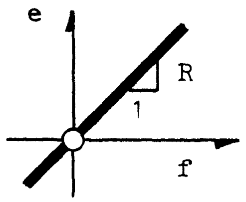
$$f = \Phi_{fe}(e) = \mathbb{G}(e)$$



These clearly reduce to the ordinary linear relations in the special case, namely:

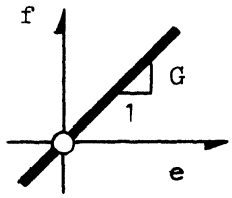
LINEAR RESISTANCE

$$e = R \cdot f$$



LINEAR CONDUCTANCE

$$f = G \cdot e$$



Whenever steady-state resistance (or conductance) is present in any one-port, it is clear from energy continuity that available energy is dissipated into heat in the amount:

$$\mathcal{P}_d = f \cdot \mathbb{R}(f) \qquad \vdots \qquad \mathcal{P}_d = e \cdot \mathbb{G}(e)$$

since no energy is stored in a purely resistive element. For the linear case:

$$\mathcal{P}_d = Rf^2 \qquad \vdots \qquad \mathcal{P}_d = Ge^2$$

which is well-known.

Let us now consider that we have a one-port system containing "ℓ" separate resistances. Then the total dissipation must be:

$$\mathcal{P}_d = \sum_{j=1}^{\ell} \mathcal{P}_{dj} = \sum_{j=1}^{\ell} e_j \cdot f_j$$

$$\mathcal{P}_d = \mathbf{f} \cdot \mathbb{R} \qquad \vdots \qquad \mathcal{P}_d = \mathbf{e} \cdot \mathbb{G}$$

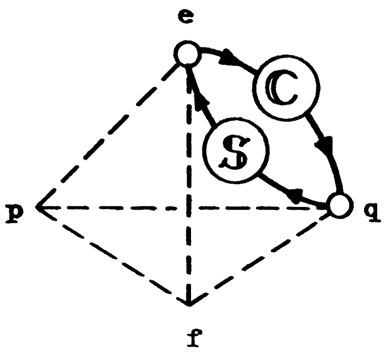
if $\mathbf{e}$, $\mathbf{f}$, $\mathbb{R}$, and $\mathbb{G}$ are treated as ℓ- coordinate vectors. Thus the energy dissipation is a scalar additive function, summed over all available energy sinks.

Moreover, if we define the functions:

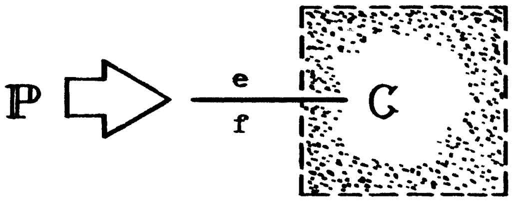
CONTENT: $P_f \equiv \int^f \mathbb{R}(f)df$ $\vdots$ COCONTENT: $P_e \equiv \int^e \mathbb{G}(e)de$

Then it is clear that the dissipation $\mathcal{P}_d = P_f + P_e$ and we obtain a dualistic generalization of the RAYLEIGH dissipation function, which we shall investigat later. This particular choice of terminology follows from the work of CHERRY, and MILLAR, both separately, and together. However, very similar notions ste from the "power function" concept of WELLS and FJERTES.

2. Capacitance Relations and Generalized Potential Energy



State Relations



Potential Energy Storage

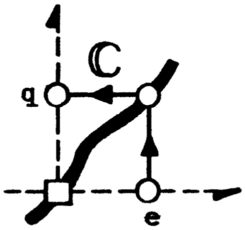
$$\boxed{\mathbb{P}(t) \equiv d\mathbb{E}_p/dt}$$

For any one-port element, the generally nonlinear static relationship between effort and displacement can always be considered as a capacitance relationship. Again we distinguish between two converse relations, namely:

$$e \rightarrow \mathbb{C} \rightarrow q$$

GENERALIZED
DISPLACEANCE

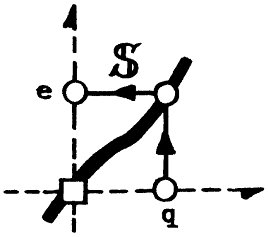
$$q = \Phi_{qe}(e) = \mathbb{C}(e);$$



$$q \rightarrow \mathbb{S} \rightarrow e$$

GENERALIZED
EFFORTANCE

$$e = \Phi_{eq}(q) = \mathbb{S}(q);$$



These clearly reduce to the ordinary <u>linear</u> relations in the special case, namely:

LINEAR CAPACITANCE	LINEAR SUSCEPTANCE
$q = C \cdot e$	$e = S \cdot q$

Whenever capacitance is present in any one-port, it is clear from energy continuity that available energy is <u>stored</u> as <u>generalized potential energy</u> in the amount:

$$E_p \equiv 2U \equiv e \cdot q$$

$E_p = e \cdot \mathbb{C}(e)$	$E_p = q \cdot \mathbb{S}(q)$

since no energy is dissipated in a (purely) capacitive element. For the linear case:

$U = (1/2)\ Ce^2$	$U = (1/2)\ Sq^2$

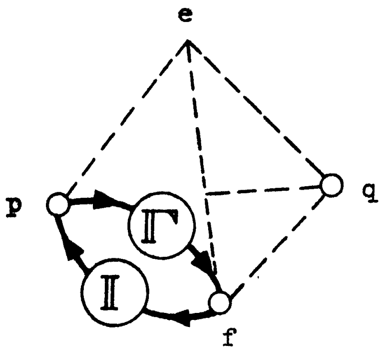
which is well-known.

For the general case we shall find it useful to define the dual pair of energy functions, where $2U = U_e + U_q$:

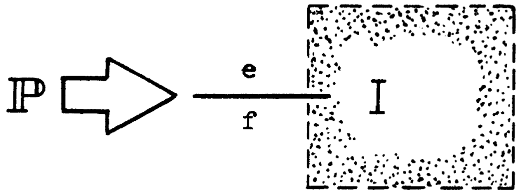
COPOTENTIAL ENERGY	POTENTIAL ENERGY
$U_e \equiv \int^e \mathbb{C}(e) \cdot de$	$U_q \equiv \int^q \mathbb{S}(q) \cdot dq$

which will later permit us to keep proper energy books. The first concepts of dualistic or complementary energy forms date back to the work of CLERK MAXWELL and ENGESSER, but these ideas are fully developed in the work of CHERRY and MILLAR previously cited.

3. Inertance Relations and Generalized Kinetic Energy



State Relations



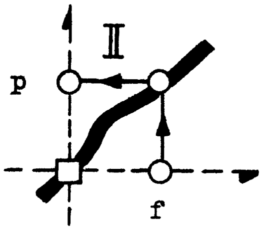
$$\boxed{\mathbb{P}(t) \equiv d\mathbb{E}_k/dt}$$

For any one-port element the generally nonlinear static relationship between flow and momentum can always be considered as an inertance relationship. Once again, we distinguish the two converse aspects, namely:

$$f \rightarrow \mathbb{I} \rightarrow p$$

GENERALIZED
MOMENTANCE

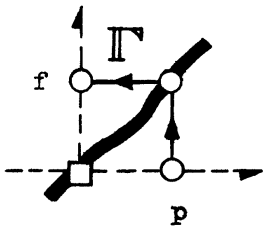
$$p = \Phi_{pf}(f) = \mathbb{I}(f);$$



$$p \rightarrow \Gamma \rightarrow f$$

GENERALIZED
FLUANCE

$$f = \Phi_{fp}(p) = \Gamma(p);$$



These clearly reduce to the ordinary linear relations in the specia case, namely:

LINEARIZED
INERTANCE

$$p = I \cdot f$$

LINEARIZED
FLUANCE

$$f = \Gamma \cdot p$$

Whenever inertance is present in any one-port, it follows from energy continuity that available energy is stored as generalized kinetic energy in the amount:

$$E_k \equiv 2T \equiv f \cdot p$$

$$E_k = f \cdot \mathbb{I}\,(f) \qquad\qquad E_k = p \cdot \Gamma\,(p)$$

since no energy is dissipated in a purely inertive element. For the linear case:

$$T = (1/2)\; I\, f^2 \qquad\qquad T = (1/2)\; \Gamma\, p^2$$

which is well-known.

As before, for the general nonlinear case, we shall find it useful to define the dual pair of kinetic energy functions:

COKINETIC
ENERGY

$$T_f = \int^f \mathbb{I}\,(f) \cdot df$$

KINETIC
ENERGY

$$T_p = \int^p \Gamma\,(p) \cdot dp$$

which are needed for the energy principles to be discussed. As before, it follows that $2T = T_f + T_p$.

G. The Three State-Determined Elements (R, C, I)

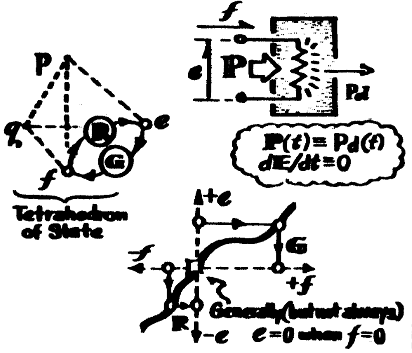
We see from the above that all primitive state-determined relations can be expressed by the three 1-port elements:

R —— ; C —— ; I ——

The above properties of these elements may all be compactly summarized in the single grand tabulation attached. The linearized values of the parameters (R, C, I) are indicated; in the general case the corresponding nonlinear relations must be generated by algebraic or by hyperpolydral functions.

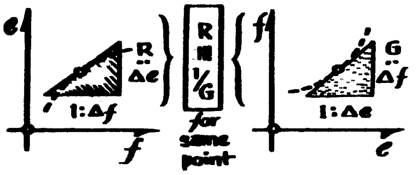
Resistance – Conductance Relations

RESISTANCE RELATION: $e = \Phi_R(f)$
Converse $\mathbb{R}$'s
CONDUCTANCE RELATION: $f = \Phi_G(e)$



$f \to \boxed{\Phi_R} \to e$ } CAUSAL BLOCK DIAGRAMS { $e \to \boxed{\Phi_G} \to f$

Local Resistance R — Local Conductance G

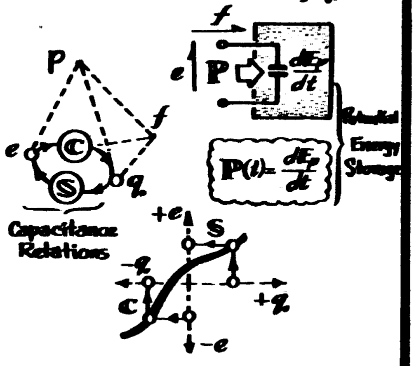


$$de = \underbrace{\frac{\partial\Phi_R}{\partial f}}_{=} \cdot df$$
$$de = R \cdot df$$
$$R \equiv de/df$$

$$df = \underbrace{\frac{\partial\Phi_G}{\partial e}}_{=} \cdot de$$
$$df = G \cdot de$$
$$G \equiv df/de$$

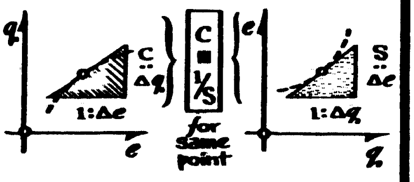
Capacitance Relations

DISPLACEANCE RELATION: $q = \Phi_C(e)$
Converse $\mathbb{R}$'s
EFFORTANCE RELATION: $e = \Phi_S(q)$



$e \to \boxed{\Phi_C} \to q$ } CAUSAL BLOCK DIAGRAMS { $q \to \boxed{\Phi_S} \to e$

Local ⟨ Displaceance / CAPACITANCE ⟩ C — Local Effortance S

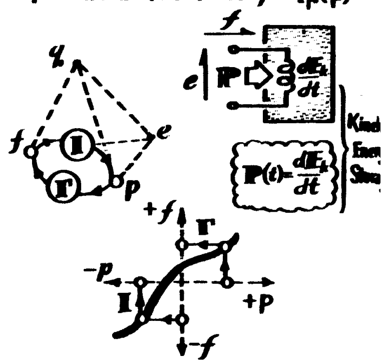


$$dq = \underbrace{\frac{\partial\Phi_C}{\partial e}}_{=} \cdot de$$
$$dq = C \cdot de$$
$$C \equiv dq/de$$

$$de = \underbrace{\frac{\partial\Phi_S}{\partial q}}_{=} \cdot dq$$
$$de = S \cdot dq$$
$$S \equiv de/dq$$

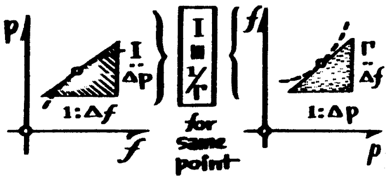
Inertance Relations

MOMENTANCE RELATION: $p = \Phi_I(f)$
Converse $\mathbb{R}$'s
FLUANCE RELATION: $f = \Phi_\Gamma(p)$



$f \to \boxed{\Phi_I} \to p$ } CAUSAL BLOCK DIAGRAMS { $p \to \boxed{\Phi_\Gamma} \to f$

Local ⟨ Momentance / INERTANCE ⟩ I — Local Fluance Γ



$$dp = \underbrace{\frac{\partial\Phi_I}{\partial f}}_{=} \cdot df$$
$$dp = I \cdot df$$
$$I \equiv dp/df$$

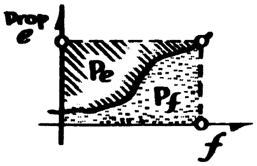
$$df = \underbrace{\frac{\partial\Phi_\Gamma}{\partial p}}_{=} \cdot dp$$
$$df = \Gamma \cdot dp$$
$$\Gamma \equiv df/dp$$

ENERGY DISSIPATION: Resistance Loss

Generalized Energy Loss }
$$d(ef) = edf + fde$$
$$2P \equiv ef = \underbrace{\int^f edf} + \underbrace{\int^e fde}$$
$$\text{or } 2P = P_f + P_e$$

Content $P_f \equiv \int^f edf = P(f)$

Cocontent $P_e \equiv \int^e fde = P(e)$



LINEAR CASE: Constant Resistance

$e = Rf;\ f = \frac{1}{R} \cdot e$; $R \equiv$ RESISTANCE

$$P_f \equiv \int^f edf = \int_0^f Rfdf = \frac{R}{2} f_2$$
$$P_e \equiv \int^e fde = \int_0^e \frac{1}{R}\, ede = \frac{1}{2R} e^2$$
$$2P \equiv P_f + P_e = \frac{R}{2} f^2 + \frac{1}{2R} e^2 = Rf^2 = \frac{e^2}{R}$$

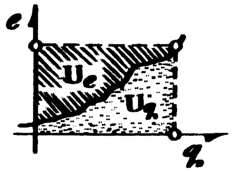
∴ For Linear Case Only $\boxed{P_f = P_e = P}$

POTENTIAL ENERGY: Capacitive Storage

Generalized Potential Energy }
$$d(eq) = edq + qde$$
$$2U \equiv eq = \underbrace{\int^q edq} + \underbrace{\int^e qde}$$
$$\text{or } 2U = U_q + U_e$$

(Normal) Potential Energy: $U_q \equiv \int^q edq = U(q)$

Complementary P.E. or Copotential Energy: $U_e \equiv \int^e qde = U(e)$



LINEAR CASE: Constant Capacitance

$q = Ce$; $e = \frac{1}{C} q$; $C \equiv$ CAPACITANCE

$$U_q \equiv \int^q edq = \int^q \frac{1}{C} q dq = \frac{1}{2C} \cdot q^2$$
$$U_e \equiv \int^e qde = \int^e cede = \frac{C}{2} \cdot e^2$$
$$2U \equiv U_q + U_e = \frac{1}{2C} q^2 + \frac{C}{2} e^2 = \frac{q^2}{C} = ce^2$$

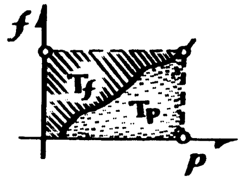
∴ For Linear Case Only $\boxed{U_q = U_e = U}$

KINETIC ENERGY: Inertive Storage

Generalized Kinetic Energy }
$$d(fp) = fdp + pdf$$
$$2T \equiv fp = \underbrace{\int^p fdp} + \underbrace{\int^f pdf}$$
$$\text{or } 2T = T_p + T_f$$

(Normal) Kinetic Energy: $T_p \equiv \int^p fdp = T(p)$

Complementary K.E. or Cokinetic Energy: $T_f \equiv \int^f pdf = T(f)$



LINEAR CASE: Constant Inertance

$p = If$; $f = \frac{1}{I} \cdot p$; $I \equiv$ INERTANCE

$$T_p \equiv \int^p fdp = \int^p \frac{1}{I} pdp = \frac{1}{2I} \cdot p^2$$
$$T_f \equiv \int^f pdf = \int^f I f df = \frac{I}{2} \cdot f^2$$
$$2T \equiv T_p + T_f = \frac{1}{2I} p^2 + \frac{I}{2} f^2 = \frac{p^2}{I} = I$$

∴ For Linear Case Only $\boxed{T_p = T_f = T}$

Background Reading -- State-Determined Systems

(1) MAXWELL, James Clerk. Matter and Motion

Starting from elementary concepts Maxwell demonstrates in a beautiful fashion, and without recourse to any sophisticated mathematical formulation, his various views on mechanics and dynamics.

(2) HERTZ, Heinrich. The Principles of Mechanics

Much is to be gained merely by reading Hertz's introduction to this short work. One is lead by his reasoning to a deeper insight into the fundamentals of mechanics. He points out, in particular, the ambiguity in Newton's definition of force, as it is implied by the three laws of motion.

(3) MAGIE, William Francis. A Source Book in Physics

Material of historical interest is presented on most of the contributors to physical science. In particular, chapters are devoted to the work of Maxwell and Hertz.

(4) LANCZOS, Cornelius. The Variational Principles of Mechanics

Again, the introduction serves as an excellent appreciation of the basis of the variational approach to the description of dynamical situations. In particular, the fundamental differences between Newtonian dynamics and the energetic method of Euler and Lagrange is pointed out; namely, the former views a system as characterized by the momenta of, and the force interactions among, its elements, while the latter relies on constraints upon the potential and kinetic energies.

(5) ROUTH, Edward John. Dynamics of a System of Rigid Bodies (First edition, 1868)

(6) WHITTAKER, E. T. Analytical Dynamics (First edition, 1904)

(7) WEBSTER, Arthur Gordon. The Dynamics of Particles (First edition, 1904)

These are standard classical references on dynamics. They represent specific excellent integrations of Newtonian and variational mechanics.

However, the applications are largely restricted to problems of academic interest.

(8) BIRKHOFF, George D. Dynamical Systems (Published in 1927)

This is a classical modern treatise on the dynamics of state-determined systems which developed out of a series of lectures given by Birkhoff in 1920.

(9) CHERRY, E. Colin. Duality, Partial Duality and Contact Transformations

The application of Hamiltonian mechanics to electro-magnetic systems is indicated.

Background Reading--One-Port State Determined Elements

(1) CHERRY, E. C. Generalized Concepts of Networks, Proceedings of the Symposium on Information Networks, Polytechnic Institute of Brooklyn, New York, 1954, pp. 176-177.

(2) CHERRY, E. C. Some General Theorems for Non-linear Systems Possessing Reactance, Phil. Mag. (7), v 42, p. 1161 (1951).

(3) CHERRY, E. C. The duality between interlinked electric and magnetic circuits and the formation of transformer equivalent circuits, Proc. Phys. Soc. B, v 62, p. 101 (1949).

(4) CHERRY, E. C. and W. MILLAR, Some New Concepts and Theorems Concerning Nonlinear Systems, in Automatic and Manual Control, Butterworths, London, 1952, pp. 263-274.

(5) MILLAR, W. Some General Theorems for Non-linear Systems Possessing Resistance, Phil. Mag. (7), v 42, p. 1150 (1951).

(6) WELLS, D. A. J. App. Phys., v 16, p. 535 (1945).

(7) FUERTES, F. A. On the Power Function, J. App. Phys., v 17, p. 712 (1946).

ground Reading -- Models and Analogs

ROSENBLUETH and Norbert WIENER. The Role of Models in Science, *Philosophy of Science*, Vol. 12, No. 4 (October, 1945) pp. 316-321.

ARBER, Agnes. Analogy in the History of Science in *Studies and Essays in the History of Science and Learning*, Schuman, New York, 1944, pp. 221-233.

ZINSSER, Hans H., M.D. Pitfalls of Physiological Modelling, *University of Southern California Medical Bulletin*, pp. 6-13 (July, 1953).

BRODBECK, May. Models, Meaning and Theories in *Decisions, Values and Groups*, Vol. I, (1960).

DEUTSCH, Karl W. Mechanism, Organism, and Society: Some Models in Natural and Social Science, *Philosophy of Science*, Vol. 18, No. 3, July, 1951, pp. 230-252.

JONES, Richard W. Models, Analogues and Homologues in *Regelungstechnik: Moderne Theorien und ihre Verwendbarkeit*, pp. 326-328.

SCHOENFELD, J. C. Analogy of Hydraulic, Mechanical, Acoustic and Electric Systems, *Applied Scientific Research*, Section B, Vol. 3, pp. 417-450.

H. The Concept of Circuits and Networks

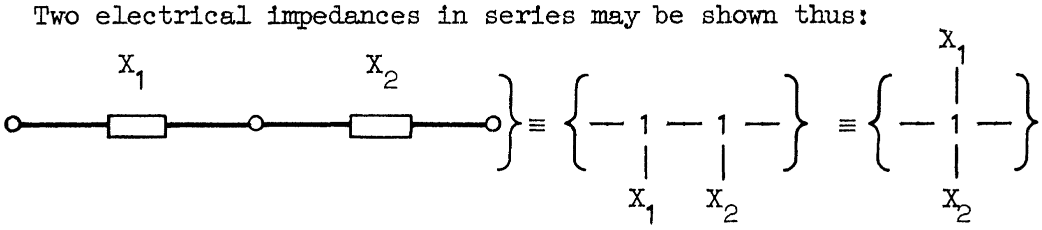
Classical dynamics has been primarily--perhaps nearly exclusively--concerned with reticular systems and processes which can be effectively conceived as composed of state-determined one-ports suitably interconnected. These models generally assume storage and dissipation of energy at a finite number of localized regions, "lumps", or "points": e.g., "mass points" in mechanics; "lumped circuits" in electricity. Such substitutes for the actu underlying field continuum have often been remarkably useful and dramaticall productive. The relations between the macroproperties of the one-port lump impedances and the microproperties of the continuous fields we shall treat the next article. Here we shall be concerned with certain of the system pr erties of lumped constant systems.

In terms of the previous relationships it is now possible to general the classical "mechanical system of many particles" or the traditional "ele trical network" to deal with any engineering system in which the "m" primitive parts are all state-determined one-port elements in any medium, each containing generally nonlinear resistance, capacitance, and inertance properties. It is not the least necessary that the system be differentiated in regard to heterogeneity of medium, since the basic relations above are vali for all media.

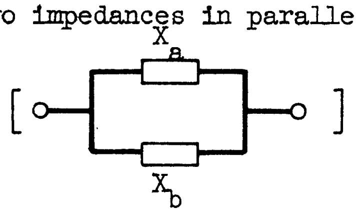
The Structure of Networks

It is first necessary to show that all possible structural combinations of one-ports may be obtained using only the two ideal junctions (0, 1 Electrical symbolism offers the most efficient explanatory language, but we can readily verify the result for mechanical systems.

Two electrical impedances in series may be shown thus:

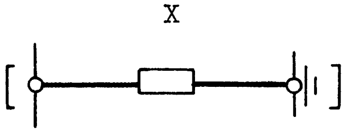


While two impedances in parallel are written



$$\equiv \begin{array}{cccccc} & & & X_a & & \\ & & & | & & \\ [\,— & 0 & — & 0 & — &] \\ & | & & & & \\ & X_b & & & & \end{array} \equiv \begin{array}{cccc} & X_a & & \\ & | & & \\ [\,— & 0 & — &] \\ & | & & \\ & X_b & & \end{array}$$

Any "grounded" impedance (i.e., where one of the efforts is the zero potential) can always be written:

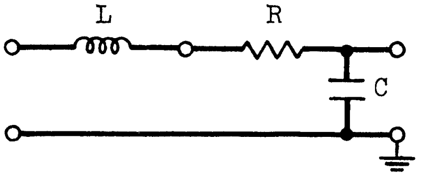


$$\equiv \begin{array}{cccccc} & & & X & & \\ & | & & | & & \\ [& 0 & — & 1 & — & E] \\ & | & & & & \| \\ & & & & & 0 \end{array} \equiv \begin{array}{cccc} & | & & \\ [& 0 & — & X\ ] \\ & | & & \end{array}$$

since the one-junction can be absorbed directly into the impedance relationship, itself. Similar results hold for the dual situation, which is particularly significant for mechanical inertias, namely:

$$\begin{array}{ccccccc} & | & & & & & \\ [& 1 & — & 0 & — & \begin{array}{c}\text{Zero}\\ \text{Flow}\end{array} &] \\ & | & & | & & & \\ & & & X & & & \end{array} \equiv \begin{array}{cccc} & | & & \\ [& 1 & — & X\ ] \\ & | & & \end{array}$$

With these properties and conventions, the damped electrical oscillator:



could be written as the radical:

$$\begin{array}{ccccc} & R & & & \\ & \bullet & & & \\ — & 1 & \bullet & 0 & — \\ & \bullet & & \bullet & \\ & L & & C & \end{array}$$

Using the facts ($L \subseteq I$, $R \subseteq R$, $C \subseteq C$), the generalized diagram would give:

$$\begin{array}{ccccc} & R & & & \\ & \bullet & & & \\ — & 1 & \bullet & 0 & — \\ & \bullet & & \bullet & \\ & I & & C & \end{array}$$

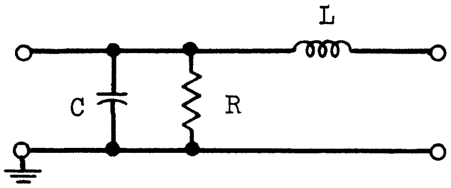
The electrical dual structure may be drawn immediately by employing the relations:

$$1 \rightleftharpoons 0 \quad ; \quad R \rightleftharpoons R \quad ; \quad I \rightleftharpoons C$$

This results in the system:

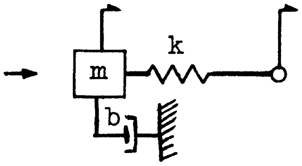
$$\begin{array}{ccccc} & R & & & \\ & \cdot & & & \\ — & 0 & \cdot & 1 & — \\ & \cdot & & \cdot & \\ & C & & I & \end{array}$$

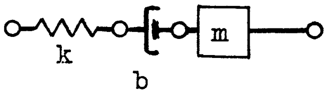
which would be sketched conventionally as:



On the other hand, both direct and dual systems may be diagramed for mechanical systems, namely:

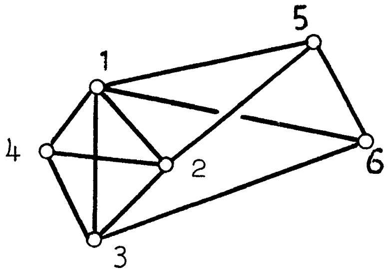
DIRECT	DUAL
$\begin{array}{ccccc} & R & & & \\ & \cdot & & & \\ — & 1 & — & 0 & — \\ & \cdot & & \cdot & \\ & I & & C & \end{array}$	$\begin{array}{ccccc} & R & & & \\ & \cdot & & & \\ — & 0 & — & 1 & — \\ & \cdot & & \cdot & \\ & C & & I & \end{array}$





On the General Significance of Networks

Thus in practical systems, we are not usually confronted with a single one-port element, but rather with a plurality of one-ports interconnected through ideal energy functions. In electrical science such networks are customarily represented by a meshwork of lines, each one of which represents a general one-port impedance element:



Moreover each such element may itself be a complex net and so on, *ad infinitum*, but this is immaterial. The points are usually called terminals or nodes. The line joining any two nodes is called a leg or branch, and any closed path made up of branches is called a loop or circuit or mesh. The topological properties of such networks are then obvious from our previous treatment.

Frequently, it has been assumed that all networks can be constructed from one-port elements. This is of course not true, and it is possible to construct other kinds of networks which contain multiports of various kinds; even since the earliest days of electromagnetism, following the work of Joseph HENRY and Michael FARADAY, the role of mutual induction--a two-port phenomenon-- has been of signal importance.

But this generalized network concept is not limited to electrical science alone. Largely through the pioneering work of Gabriel KRON, the true role played by reticular fields and generalized nets is now better understood. References are given in the reading list to applications in such diverse fields as:

NUMERICAL ANALYSIS
DIRICHLET PROBLEM
FIELD PROBLEMS
SCHRODINGER EQUATION
ELASTICITY and PLASTICITY
CONFORMAL MAPPING
NONLINEAR NETS
ALGEBRAIC TOPOLOGY
RADIATION ANALYSIS
FLUID FLOW

Background Reading - Networks

PHILLIPS, H. B., and N. WIENER, Nets and the Dirichlet Problem, *Journal of Mathematics and Physics*, Vol. 2, pp. 105-124 (1923).

LYUSTERNIK, L. A. On Electrical Modelling of Symmetric Matrices, *Uspekhi Matem.* Nauk (N. S.) Vol. 41, pp. 198-200 (1949).

VON MISES, R. On Network Methods in Conformal Mapping and in Related Problems. Construction and Application of Conformal Maps, U. S. Dept. of Commerce, Natl. Bur. of Standards, *Applied Math. Ser.* Vol. 18, pp. 1-6 (1952).

BIRKHOFF, G. D., and J. B. DIAZ. Non-linear Network Problems, *Quarterly of Applied Mathematics*, Vol. 13, No. 4, pp. 431-443 (1956

OPPENHEIM, A. K. Radiation Analysis by the Network Method, *Transactions of the ASME*, Vol. 78, pp. 75-735 (1956).

KRON, G. Numerical Solution of Ordinary and Partial Differential Equatio by means of Equivalent Circuits, *Journal of Applied Mecha ics*, Vol. 16, pp. 176-186 (1945).

-----. Equivalent Circuit of the Field Equations of Maxwell, *Proceedings of the IRE*, Vol. 32, pp. 289-299 (1944).

-----. Electric Circuit Models of the Schroedinger Equation, *Physics Review*, Vol. 67, pp. 39-43 (1945).

-----. Equivalent Circuits of the Elastic Field, *Journal of Applied Mechanics*, Vol. 11, pp. A149-161 (1944).

-----. Equivalent Circuits of Compressible and Incompressible Fluid Flow Fields, *Journal of Aeronautical Science*, Vol. 12, pp. 221-231 (1945).

-----. A Set of Principles to Interconnect the Solutions of Physical Systems, *Journal of Applied Physics*, Vol. 24, pp. 965-980 (1953).

-----. Solution of Complex Non-linear Plastic Structures by the Method of Tearing, *Journal of Aeronautical Science*, Vol. 23, pp. 557-562 (1956).

BRANIN, F. H. Kron's Method of Tearing and its Applications, *Proceedings of the Second Midwest Symposium on Circuit Theory*, Michig State University, pp. 2.1-2.79 (1956).

ROTH, J. P. An Application of Algebraic Topology: Kron's Method of Tearing, *Quarterly of Applied Mathematics*, Vol. XVII, No. 1, p. 1 (April, 1959).

XIII. Distribution of Energy over Space, Time and Frequency

A. Introduction

We are concerned in this section with a quantitative description of the distribution of power and energy over space and time. The spatial distribution requires that we consider the properties of continuous and reticular fields. In the general case, these fields are nonstationary or unsteady. but for most engineering purposes we may consider the fields as pseudo-static (or quasi-stationary). Such fields for energetic systems are governed by scalar potential functions, together with their derived and associated vector fields. If the fields are truly dynamic, we can preserve the field language if we employ retarded potentials and a corresponding integral formulation.

At any given point (or bond) in an energetic field, the local power state is instantaneously related to the boundary or environmental power states through dynamic transfer characteristics or operators. These may be employed either to describe the local behavior in the time domain or to interpret the response characteristics in terms of frequency or spectral sensitivity. Thus the local energy distribution over time of any linear (or linearizable) system may alternatively be viewed as a distribution of the same energy over the frequency band.

These fundamental facts and concepts, sufficient to deal with multiport systems, are few in number and are outlined in the paragraphs below.

In Part IV we treated the continuity of energy in a generalized field while Part V introduced the concept of field reticulation and the corresponding reticulation of field energy and power. These reticular energies were then evaluated for state-determined elements. We now proceed to the precise restatement of the reticular energy principles and power balances for state-determined systems. This is the form in which the generalized energy concepts were first obtained; but, consistent with the assumptions of modern relativistic mechanics and quantum physics, it is energy and not material structure which is the foundation point for rational science.

B. Energy Principles for State-Determined Systems

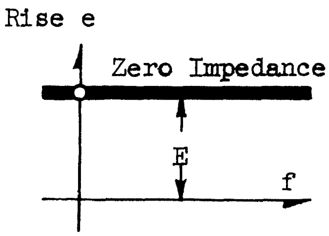
Boundary Conditions as Sources of Flow and Effort

It is useful to introduce the classically important concepts of pure sources of effort and flow as indicated below. These correspond, for the continuous case, to the traditional DIRICHLET and NEUMANN boundary conditions. Moreover, it is possible to represent nearly all boundary tra ports of energy in terms of the generalized HELMHOLTZ and THEVENIN equivalent sources, the names corresponding to the traditional linear electrical equivalents. Merely by extending the boundaries of the system to include the source impedance functionals, we can represent all power transport across the boundaries of an n-port state-determined system in terms of the ideal DIRICHLET or NEUMANN sources. In that which follows below, for simplicity, let us assume all boundary ports as equivalent to a finite or infinite set of such ideal elements.

DIRICHLET PORT: $\boxed{\text{Element: } E \text{ —}}$

CONSTANT EFFORT SOURCE:

$$\begin{cases} e = E = \text{constant} \\ \dfrac{\partial e}{\partial f} \equiv 0 \end{cases}$$



$$\text{Content } P_f \equiv \int^f (\text{Drop } e)\,df = \int^f (-E)\,df$$

or $\boxed{P_f = -\,Ef}$

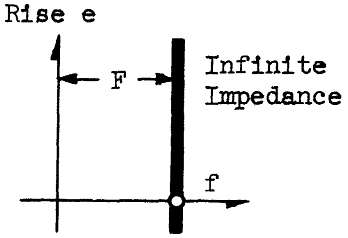
(i.e., Power GAIN = E • f)

Cocontent $\boxed{P_e \equiv 0}$

NEUMAN PORT: $\boxed{\text{Element: } F \text{ —}}$

CONSTANT FLOW SOURCE:

$$\begin{cases} f = F = \text{constant} \\ \dfrac{\partial f}{\partial e} \equiv 0 \end{cases}$$



Content $\boxed{P_f \equiv 0}$

$$\text{Cocontent } P_e \equiv \int^e f\,de = \int^e F(-$$

or $\boxed{P_e = -\,Fe}$

(i.e., Power GAIN = F • e)

Pure Sources of Effort and Flow

Generalized Dynamical Relationships

The classical LAGRANGIAN function was historically the first introduced to deal with generalized state-determined systems. This may be defined in terms of our relations above in the form:

$$\text{Lagrangian } \mathbb{L} = \mathbb{L}(f,q) = \mathbb{T}_f - \mathbb{U}_q$$
$$= \sum \text{Cokinetic Energies} - \sum \text{Potential Energies}$$
$$= \text{Total Free Energy}$$

that is:

$$\mathbb{T}_f = \sum_{i=1}^{m} T_{fi}(f_i) = T_{f1} + T_{f2} + \cdots + T_{fm}$$

$$\mathbb{U}_q = \sum_{i=1}^{m} U_{qi}(q_i) = U_{q1} + U_{q2} + \cdots + U_{qm}$$

We may then write the Lagrange Equation in the form:

$$\boxed{\frac{d}{dt}\left(\frac{\partial \mathbb{L}}{\partial \boldsymbol{f}}\right) - \frac{\partial \mathbb{L}}{\partial \mathbf{q}} + \frac{\partial \mathbb{P}_f}{\partial \boldsymbol{f}} = 0}$$

since we can arrange that $\mathbb{P}_f$ includes all the energy sources as well as sinks.

However, it was not realized until comparatively recently that there also exists a completely parallel and dual form of the Lagrange Equation expressible entirely in terms of the effort vector, $\mathbb{e}$, and momentum vector, $\mathbb{p}$. Here, the dual Lagrangian would be a complementary free energy in terms of the total copotential energy less the total kinetic energy. Thus we may express the two dual energy formulations side-by-side in the form of Table I below.

TABLE I

LAGRANGE EQUATIONS

Classical Form	Dual Form
$\frac{d}{dt}\left(\frac{\partial \mathbb{T}_f}{\partial \mathscr{f}}\right) + \frac{\partial \mathbb{U}_q}{\partial \mathbb{q}} + \frac{\partial \mathbb{P}_f}{\partial \mathscr{f}} = 0$	$\frac{d}{dt}\left(\frac{\partial \mathbb{U}_e}{\partial \mathbb{e}}\right) + \frac{\partial \mathbb{T}_p}{\partial \mathbb{p}} + \frac{\partial \mathbb{P}_e}{\partial \mathbb{e}} = 0$

COKINETIC ENERGY: $\mathbb{T}_f \leftrightarrow \mathbb{U}_e$: COPOTENTIAL ENERGY

POTENTIAL ENERGY: $\mathbb{U}_q \leftrightarrow \mathbb{T}_p$: KINETIC ENERGY

CONTENT: $\mathbb{P}_f \leftrightarrow \mathbb{P}_e$: COCONTENT

Let us next see how we may obtain a generalized power balance in any state-determined system. The dynamics of a normal Lagrangian system would be governed by the pair of equations:

$$\frac{d}{dt}\left(\frac{\partial \mathbb{T}_f}{\partial f_i}\right) + \frac{\partial \mathbb{U}_q}{\partial q_i} + \frac{\partial \mathbb{P}_f}{\partial f_i} = 0$$

$$\frac{d}{dt}(q_i) = f_i$$

If these expressions are multiplied together, and summed, there results:

$$\boxed{\frac{d}{dt}(\mathbb{T}_p + \mathbb{U}_q) + P_d = P_a}$$

where P_a is the active content (or power supplied to the system) from all energy sources. By considering these sources as energy ports, we obtain the previous energy expression:

$$(d\mathbb{E}/dt) + \mathcal{P}_d = \mathbb{P}$$

Alternatively, we may carry out a power balance in terms of the system <u>Hamiltonian</u> function

$$\mathbb{H} \equiv \mathscr{f} \cdot \mathbb{p} - \mathbb{L}(\mathscr{f}, \mathbb{q}) = \mathbb{H}(\mathbb{p}, \mathbb{q})$$

In terms of the individual elements:

$$\mathscr{f} \cdot \mathbb{p} = \sum_i f_i p_i = \sum_i (T_{fi} + T_{pi})$$

$$\mathbb{L} = \sum_i (T_{pi} - U_{qi})$$

This gives the general result:

$$\mathbb{H} = \sum_i (T_{pi} + U_{qi}) = \mathbb{E} = \text{Total Stored Energy}$$

Then a general Hamiltonian power balance gives:

$$\boxed{\underbrace{\frac{d\mathbb{H}}{dt}}_{\substack{\text{Change of}\\ \text{Stored Energy}}} + \underbrace{\sum_i f_i \cdot \frac{\partial \mathbb{P}_{fi}}{\partial f_i}}_{\substack{\text{Energy Sources}\\ \text{and Sinks}}} = 0}$$

again merely a specialization of the reticulated energy continuity.

We may then summarize all the previous results in the form of the <u>generalized energy diagram</u> depicted in the figure below. This result follows by time-integrating the power flow for an isolated or closed system (i.e., $\sum \mathbb{P} \equiv 0$). In this case we may write:

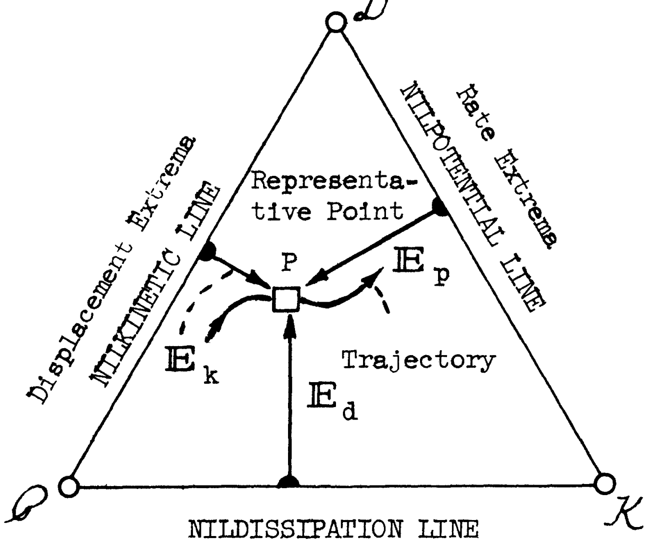
$$\mathbb{E}_k(t) \equiv \mathbb{T}_p = \sum_i T_{pi}$$

$$\mathbb{E}_p(t) \equiv \mathbb{U}_q = \sum_i U_{qi}$$

$$\mathbb{E}_d(t) \equiv \int_0^t \mathscr{P}_d(t)\,dt = \sum_i \int^t P_{di}\,dt$$

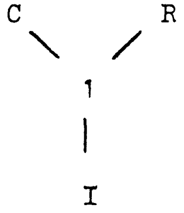
for which

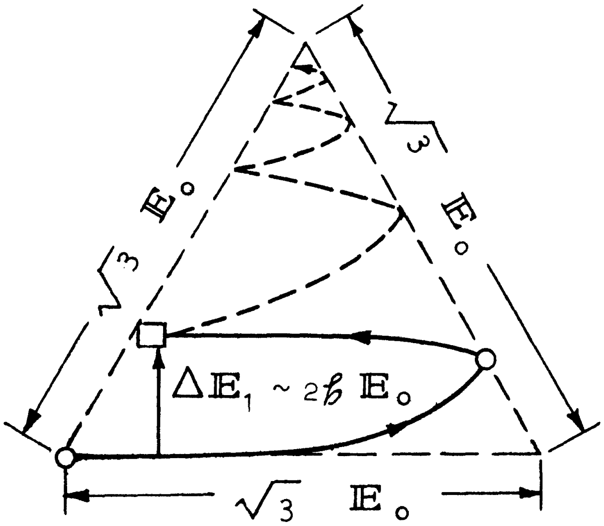
$$\boxed{\mathbb{E}_k(t) + \mathbb{E}_p(t) + \mathbb{E}_d(t) = \text{const} = \mathbb{E}_o}$$



EXAMPLE

DAMPED SECOND-ORDER SYSTEM





C. Fields, Potentials, and Transmittances

The finite macroreticulations discussed previously always represent merely approximate partitionings of the continuous field energies into a finite number of abstract individual elements of state-determined or general functional form. For the finite state case, these elements are marked by the relational symbols, $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{I}$, and therefore represent finite coordinate systems which are the natural generalizations of rigid body dynamics.

Electrical circuits and networks are but special cases in the electromagnetic domain of the general concept of reticular fields wherein the fluxes and potentials are assumed to conform to a prescribed meshwork. The lumped circuit concept is thus to electrical science what the Newtonian mass point is to mechanics and the Hookian linear spring to elasticity.

In particular, these approximations ignore more or less completely the finite velocity of energy propagation and the consequent field retardation effects. Thus, while a more rigorous field theory would formulate functional or operational relations between the field quantities, the reticular field concept presumes that simple static functions relate local or total variables.

For these and other reasons any particular system representation can at most hold only within a restricted amplitude and frequency domain. Outside these limits, we must inevitably expect discrepancies between analysis and experiment, between prediction and performance. While we can partially reduce these divergences by making the reticulation and corresponding model more complex, we cannot ever expect complete equivalence between the multiply-infinite order of physical reality and the modest finite order of our conceptual models.

Quasi-Stationary Processes and Slowly Varying Fields

Most of the problems of engineering analysis and system design li within the domain of slowly varying fields. For most cases this condition will hold whenever the dimensions of the system are small compared to the characteristic wavelengths of all disturbances. Under these circumstances the field retardation may be neglected and the field properties may be calculated as for stationary processes. In this manner static linear or nonlinear relations may be established among the principal variables which may be taken as integral forms of the corresponding field quantities .

The resultant inertance, capacitance, and resistance relations depend in addition to the material constants only upon the geometry of the field and are the result of integrations or averaging processes performed over the space coordinates. Thus only an integration with respect to time remains.

While a precise treatment of rapidly varying processes demands consideration of transient field effects, usually represented by systems of partial differential equations, a description in terms of ordinary differential equations suffices for slowly varying fields. Moreover, for the static and stationary states of equilibrium these relations reduce still further to systems of purely algebraic equations.

By contrast, the introduction of the concept of a retarded potenti for rapidly varying unsteady motion represents an attempt to preserve the field concept for the description of nonstationary phenomena in continuous media. We shall encounter a particular instance of this technique in dealing with wavelike transmission in Part XVI.

Fields and Field Analogies

The concept of a stationary field occurs in many branches of physics and engineering; for example: electrodynamics, aerodynamics, hydrodynamics, elasticity, heat conduction, and gravitational theory. Insofar as the phenomena admit the definition of scalar potential functions, their abstract form and treatment is essentially identical. This makes it possible to establish formal analogies of strict equivalence and to translate solutions and experimental results from any one field into all analogous fields. Each one of these particular media is characterized by a fundamental scalar which satisfies the Poisson or Laplace equation, by a derived field vector which is defined as the gradient of the associated scalar potential, and by an associated field vector which is tensorially related to the gradient vector. There are additional analogous concepts for each applied field which makes it only necessary to obtain the solution of a problem in one branch in detail in order to be able to predict for every other branch the similar solution merely by making the proper correspondence of terms. For our purposes here we shall find it convenient to deal entirely in terms of a generalized language as indicated in the first row of the appended table. It is of particular interest to note that associated with each scalar potential field there is, in general, an associated vector field. In keeping with our generalized symbolism, we may think of one set of fields as intrinsic or effort fields and the other set as extrinsic or flow fields. In the electrical case we have the electrostatic field and its relation with the capacitance in the form of the associated field transmittance; in the magnetostatic case, the associated magnetic flux and field permeance. It is frequently helpful and suggestive to interchange terminology and other imagery from one medium to another, so that the maximum cross-fertilization of ideas can occur. On the other hand, by manifesting that all particularizations arise from a single, common generalization, all field analogies are rigorized and a methodical approach to field concepts is made evident.

SPECIALIZATIONS OF

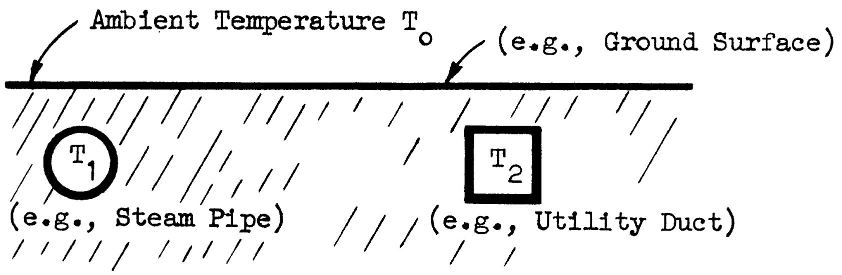
	Potential Function	Potential Difference	Equipotential Surface	Potential Gradient
GENERALIZED FIELD	u	$u_1 - u_2 = \int_1^2 \vec{U} \cdot \vec{ds}$	$u = \text{const.}$	$\vec{U} = (\pm)\vec{\nabla} u$
ELECTROMAGNETIC	$\mathcal{F}$ Magnetostatic Potential	$\mathcal{F}_1 - \mathcal{F}_2 = \int_1^2 \vec{H} \cdot \vec{ds}$	$\mathcal{F} = \text{const.}$ (Ideal Iron Surfaces)	$\vec{H} = -\vec{\nabla} \mathcal{F}$ Magnetic Field Strength
ELECTROSTATIC	$\mathcal{E}$ Electrostatic Potential	$\mathcal{E}_1 - \mathcal{E}_2 = \int_1^2 \vec{E} \cdot \vec{ds}$	$\mathcal{E} = \text{const.}$ (Ideal Conductor Surfaces)	$\vec{E} = -\vec{\nabla} \mathcal{E}$ Electric Field Strength
ELECTRIC CURRENT	$\mathcal{E}$ Electrostatic Potential	$\mathcal{E}_1 - \mathcal{E}_2 = \int_1^2 \vec{E} \cdot \vec{ds}$	$\mathcal{E} = \text{const.}$ (Electrode Surfaces)	$\vec{E} = -\vec{\nabla} \mathcal{E}$ do
STEADY TEMPERATURE	T Temperature	$T_1 - T_2 = \int_1^2 \vec{U} \cdot \vec{ds}$ Temp. Differential	$T = \text{const.}$ Isotherms	$\vec{U} = -\vec{\nabla} T$ Temperature Gradient
FLUID VELOCITY	ϕ Velocity Potential	$\phi_1 - \phi_2 = \int_1^2 \vec{V} \cdot \vec{ds}$	$\phi = \text{const.}$ Equipotential Surfaces	$\vec{V} = -\vec{\nabla} \phi$ Velocity
FLUID SEEPAGE	H Pressure Head	$H_1 - H_2 = \int_1^2 \vec{U} \cdot \vec{ds}$ Differential Head	$H = \text{const.}$ Piezometric Surfaces	$\vec{U} = -\vec{\nabla} H$ Seepage Gradient
GRAVITATIONAL	U Gravitational Potential	$U_2 - U_1 = \int_1^2 \vec{g} \cdot \vec{ds}$ Potential Difference	$U = \text{const.}$ Isogravimetric Surfaces	$\vec{g} = +\vec{\nabla} U$ Gravitational Acceleration

ISOTROPIC POTENTIAL FIELDS

Material Constant	Associated Field Vector	Associated Flux	Field Transmittance	Remarks
κ	$\vec{V} = \kappa \vec{U}$	$v = \int \vec{V} \cdot \vec{dA}$	$\mathcal{T} = v/(u_1 - u_2)$ Transmittance	$\mathcal{T} = \kappa A_n/L_n$ $= \kappa L \lambda$
μ Permeability	$\vec{B} = \mu \vec{H}$ Magnetic Induction	$\Phi = \int \vec{B} \cdot \vec{dA}$ Magnetic Flux	$\mathcal{P} = \Phi/(\mathcal{F}_1 - \mathcal{F}_2)$ Permeance	
$\mathcal{E}$ Dielectric Constant	$\vec{D} = \mathcal{E}\vec{E}$ Electric Induction	$Q = \int \vec{D} \cdot \vec{dA}$ Electric Charge	$C = Q/(\mathcal{E}_1 - \mathcal{E}_2)$ Capacitance	
σ Electric Conductivity	$\vec{J} = \sigma \vec{E}$ Current Density	$I = \int \vec{J} \cdot \vec{dA}$ Current	$G = I/(\mathcal{E}_1 - \mathcal{E}_2)$ Conductance	
k Thermal Conductivity	$\vec{J} = \overrightarrow{kU}$ Heat Flux	$Q = \int \vec{J} \cdot \vec{dA}$ Heat Flow	$G = Q/(T_1 - T_2)$ Thermal Conductance	
ρ Mass Density	$\vec{p} = \rho \vec{V}$ Momentum Flux	$W = \int \vec{p} \cdot \vec{dA}$ Mass Flow	$G = W/(\Phi_1 - \Phi_2)$	
σ Permeability	$\vec{V} = \sigma \vec{U}$ Seepage Velocity	$Q = \int \vec{V} \cdot \vec{dA}$ Seepage Flow	$F = Q/(H_1 - H_2)$ Seepage Conductance	

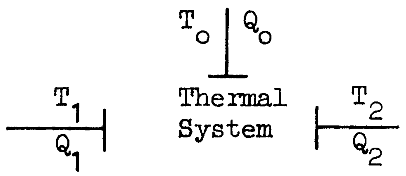
The Concept of a Field Transmittance

Consider a heat conducting medium with the configuration indicat



We might consider the situation where the temperatures were all considered as functions of time, $T_o(t)$, $T_1(t)$, $T_2(t)$, and we desire information on the resultant flow of heat, particularly, for example, the heat loss from the steam pipe or the heat flow into the utility duct.

This system may be viewed as a 3.port thermal element:



If the element can be assumed linear and in the steady state then we know that the following equivalent sets of linear equations hold:

$$\mathbb{T} = \mathbb{R} \cdot \mathbb{Q}$$

$$\begin{bmatrix} T_1 \\ T_2 \\ T_3 \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \\ Q_3 \end{bmatrix}$$

or

$$\mathbb{Q} = \mathbb{G} \cdot \mathbb{T}$$

$$\begin{bmatrix} Q_1 \\ Q_2 \\ Q_3 \end{bmatrix} = \begin{bmatrix} G_{11} & G_{12} & G_{13} \\ G_{21} & G_{22} & G_{23} \\ G_{31} & G_{32} & G_{33} \end{bmatrix} \begin{bmatrix} T_1 \\ T_2 \\ T_3 \end{bmatrix}$$

However, since we are presumably dealing with a potential field, GREEN'S THEOREM (see below) results in MAXWELL'S RECIPROCAL RELATIONS, namely

$$R_{ij} \equiv R_{ji} \qquad ; \qquad G_{ij} \equiv G_{ji}$$

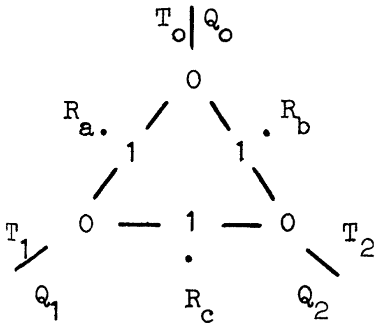
which merely express the fact that the resistance and conductance matrices are symmetric. But two additional nodicity conditions will also hold, namely:

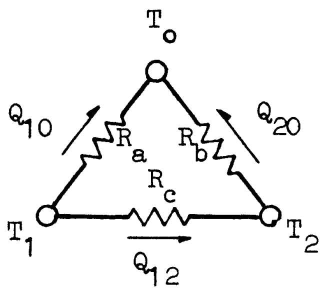
$$\text{I} \cdots \quad Q_1 + Q_2 + Q_3 \equiv 0$$

$$\text{II} \cdots \quad \mathbb{Q}\,(\mathbb{T}) \equiv \mathbb{Q}\,(\mathbb{T} + \mathbb{I}\;T) \qquad (T = \text{Const})$$

The first relation expresses the continuity condition while the second equation states the relativity condition (i.e. that the heat flow depends only upon relative, and not upon absolute, temperatures). As we discuss later in connection with trinode amplifiers, Condition I then requires that the row sums of $\mathbb{R}$ (and the column sums of $\mathbb{G}$) must vanish identically while Condition II requires that the column sums of $\mathbb{R}$ (and the row sums of $\mathbb{G}$) must also vanish.

Finally, as a result of the reciprocity and nodicity conditions above, this field problem may be conceived in terms of the following reticulation:





If the values of the three equivalent thermal resistances (R_a, R_b, R_c) were known, the instantaneous heat flows could be calculated.

This model derives its validity from the presumed superposibility (or linearity) of the governing temperature fields. This means, for instanc that we can consider the total heat loss from the steam pipe as the sum of the loss to the atmosphere (Q_{10}) and the loss to the duct (Q_{12}). While this is reasonable and obvious, it is not so obvious that these effects are ideal independent of each other, such that the loss may be assumed in the form:

$$Q_1 \equiv Q_{10} + Q_{12} = G_a(T_1 - T_o) + G_c(T_1 - T_2)$$

where G_a and G_c represent overall linear heat transmittances or thermal conductances. As we shall see, each of these conductances may be derived or estimated directly from the form (alone) of the temperature field, in the fashion:

$$\boxed{G_m = \underset{\substack{\uparrow \\ \text{Thermal} \\ \text{Conductivity}}}{k} \cdot \underset{\substack{\uparrow \\ \text{Running} \\ \text{Length}}}{L_n} \cdot \underset{\substack{\uparrow \\ \text{Field Form} \\ \text{Factor}}}{\lambda_m}}$$

Thus it is that the macroscopic properties (e.g., G_m) are related to the microscopic properties (e.g., k, λ_m) and the absolute size (e.g., L).

All the above concepts were first employed in a general and consistent fashion by James Clerk MAXWELL for electromagnetic fields. We demonstrate below that these tools have universal application and great utility. In particular, all one-port linear properties may be related very simply to the geometrical parameters (≡ size X shape) and material properties (≡ transmissivity) in terms of the overall transmittance:

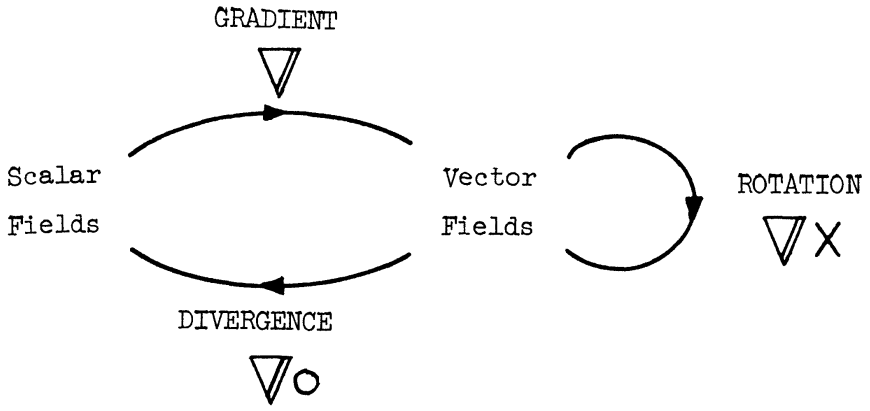
$$\boxed{(\text{Transmittance}) \equiv \underbrace{(\text{Transmissivity})}_{\text{Property}} \cdot \underbrace{(\text{Size}) \cdot (\text{Form})}_{\text{Geometry}}}$$

A Brief Survey of Vectorial Field Mechanics

It has not been customary in applied mathematics to consider the particular engineering significance of abstract concepts, nor in specific physical or engineering sciences, to display clearly the logical necessity and interrelationship of diverse, separate historical discoveries. Thus we find valuable at this point a terse summary of fundamental concepts upon which the field description of all material systems depend.

We are concerned here both with scalar quantities (u) which are the analogs of geometrical points, and also vector quantities ($\vec{U}$ or $\mathbb{U}$), the analogs of directed line segments. Pressure, density, temperature are typical scalars; while force, velocity, heat flux, current density are representative vectors. The treatment below is somewhat different than the usual in the introduction and use of matrix notation for vectoral relations.

Of course, the ordinary time and space derivatives transform scalar fields into new scalar fields, and vectors into new vectors. But in addition to these familiar operations we must consider also three additional vector derivative operations as follows:



To complete the field description, we need also to account for the material or phenomenological operators relating vectors to associated vectors in the form:

$$\mathbb{V} = \mathbb{K}\ \mathbb{U}$$

In an isotropic field, the general matrix $\mathbb{K}$ reduces to a simple scalar constant, κ.

Scalar Fields

If any scalar quantity, u, such as temperature, pressure, density, voltage, etc., is defined at every point in a region we have a scalar field. The general scalar field varies with time as well as position, which in Cartesian coordinates we would indicate by $u = u(x,y,z,t)$. But the field is clearly independent of the coordinate system in which it is measured; to emphasize this invariance it is preferable to write $u = u(\vec{R},t)$, where $\vec{R}$ measures the position vector in an arbitrary frame.

If the field is static or stationary, it is time-invariant and we have $u = u(\vec{R})$, alone.

If we connect all the adjacent points in a scalar field having equal values of u, we obtain level surfaces, isopleths, or isotimics. (These only become equipotentials if the scalar u is a potential function). In two dimensions the isotimics are curves, the analogs of the familiar contour lines. In the ensuing discussion, we shall find it convenient to depict our results, whenever appropriate, for this planar case since it is best suited to the limitations of the printed page (or blackboard!).

The First Derived Quantity: The GRADIENT of a SCALAR FIELD

This derived operation → $\boxed{\overrightarrow{\text{grad}}\; u \equiv \vec{U}}$ → Generates a VECTOR FIELD

The GRADIENT is a DERIVED VECTOR QUANTITY equal to the MAXIMUM SPATIAL RATE of CHANGE of the SCALAR FIELD FUNCTION.

The vector $\vec{U} = \vec{\nabla} u$ may be computed by matrix means; for Cartesian coordinates, this may be written:

$$U = \nabla \cdot u$$

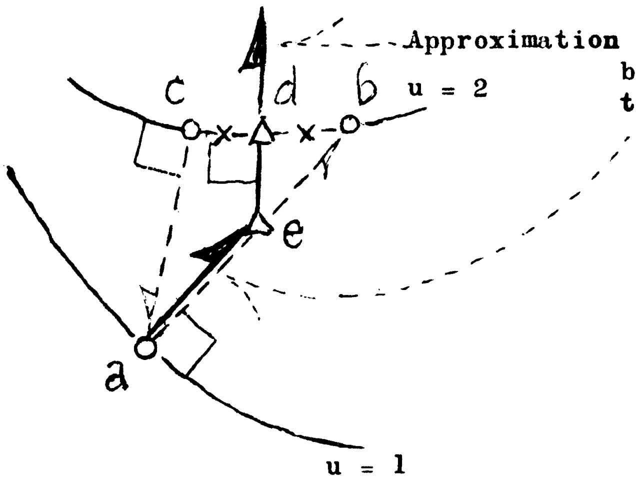
$$U = [\partial/\partial x \vdots \partial/\partial y \vdots \partial/\partial z]\; u$$

The gradient is always normal to the isotimics and measures the maximum spatial rate of change of the scalar field. In planar fields, the gradient is thus directed at right angles to the contours and measures the direction and magnitude of the steepest ascending slope.

This concept of a gradient is fundamental to the field description of physical systems. The gradient generates a vector field from the given scalar field; we might perhaps best visualize this process in terms of the gradient field derived from a set of elevation contours above a two-dimensional surface such as a plane or sphere.

In general, the gradient vector will be different at each and every point of the scalar field; we thus obtain a derived vector field which may be represented by an equivalent system of fieldlines or gradient lines.

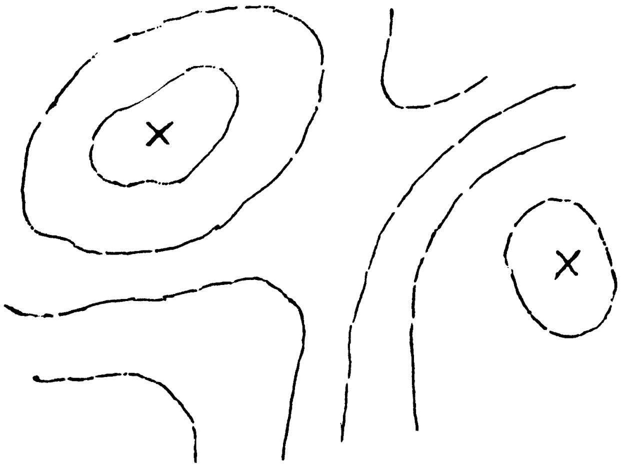
These gradient lines may be approximated by the following simple construction.



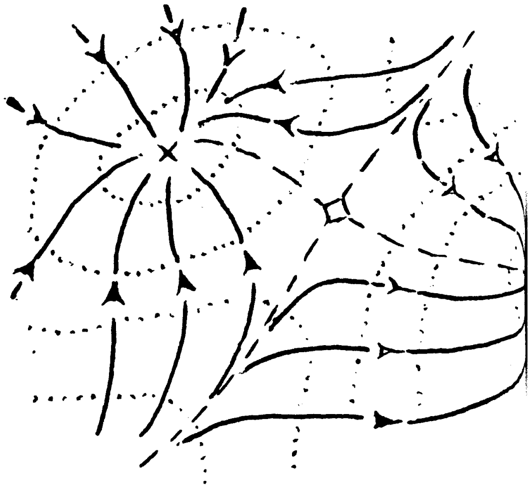
Approximation to the gradient line between u = 1 and u = 2 through point a.

Continuing in this manner until a summit is reached, an approximation to a curvilinear gradient line may be obtained. If a system of such fieldlines are drawn, any scalar field may be represented in conjugate form by the system of gradient lines as follows:

Original Representation
by
CONTOUR LINES



Conjugate Representation
by
GRADIENT LINES



Thus we have, in fact, replaced the original scalar field by an associated vector-field--the gradient field. Clearly, we could recover the original contour lines by the same process used to obtain the gradient lines.

The line integral of a gradient field between any two points is independent of the path and is equal to the difference between the values of the associated scalar at the terminal points of the path; thus

$$\int_1^2 \vec{\nabla} u \cdot d\vec{R} = u_2 - u_1$$

If the path is closed, the line integral must necessarily vanish:

$$\oint \vec{\nabla} u \cdot d\vec{R} \equiv 0$$

The significance of this last result lies particularly in the converse interpretation: namely, that

$$\text{If} \qquad \oint \vec{F} \cdot d\vec{R} \equiv 0, \text{ then } \vec{F} \equiv \overrightarrow{\text{grad}}\ u$$

where u is an associated scalar field.

The Second Derived Quantity:

(The Divergence of a Vector Field)

This derived operation $\longrightarrow$ $\boxed{\text{div} \cdot \vec{F} = \vec{\nabla} \cdot \vec{F}}$ $\longrightarrow$ Generates a SCALAR FIELD

> The DIVERGENCE is a DERIVED SCALAR QUANTITY equal to the LOCAL PRODUCTION of FIELD LINES per unit volume.

The scalar divergence $\sigma = \vec{\nabla} \cdot \vec{F}$ may be computed by matrix means; for Cartesian coordinates this inner product would be written:

$$\sigma = \nabla \circ \cdot \mathbb{F}$$

$$\sigma = [\partial/\partial x \,\vdots\, \partial/\partial y \,\vdots\, \partial/\partial z] \cdot \begin{bmatrix} F_1 \\ \hdashline F_2 \\ \hdashline F_3 \end{bmatrix}$$

The Gauss Divergence Theorem

Let us consider any closed control surface A bounding a volume in any vector field F.

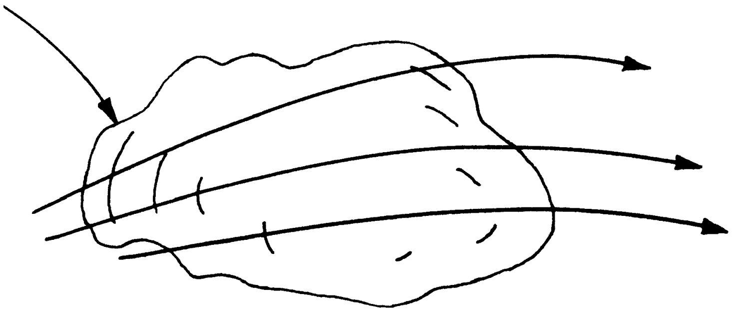
Then the following fundamental theorem due originally to Karl Friedrich GAUSS will always hold:

$$\boxed{\underbrace{\iint_A \vec{F} \cdot \vec{dA}}_{\substack{\text{Surface} \\ \text{Evaluation}}} \equiv \underbrace{\iiint_{\mathcal{V}} (\vec{\nabla} \cdot \vec{F})\, d\mathcal{V}}_{\substack{\text{Volume} \\ \text{Evaluation}}}}$$

Control Element

SURFACE: A

VOLUME : $\mathcal{V}$



The GAUSS Theorem (together with its STOKES' conjugate below) may properly be considered as the basic theorems of mathematical physics, since from them may be derived all conservation and continuity principles, as well as GREEN'S theorem and the reciprocity principle.

Field Tubes and Solenoidal Fields

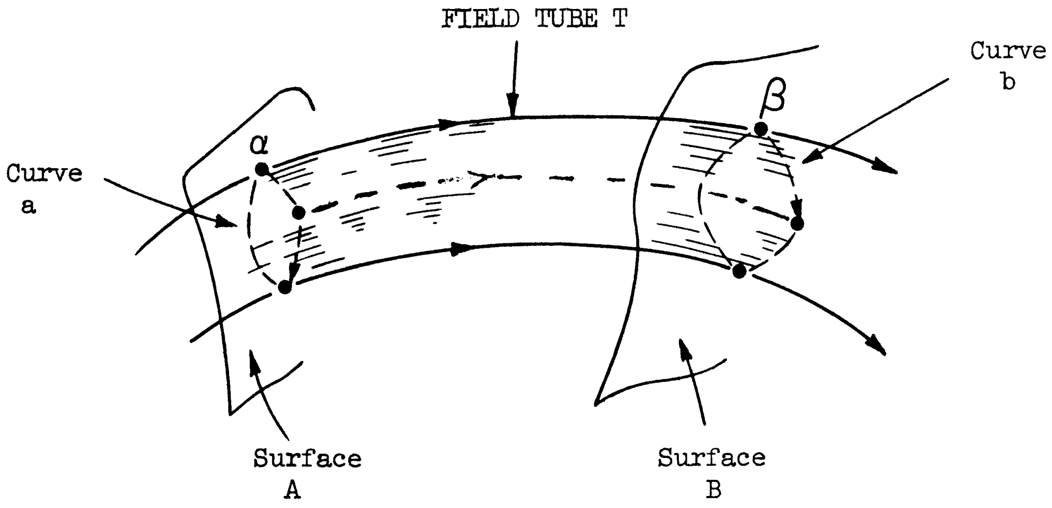
If in a general vector field, $\vec{F}$, we restrict attention to regions which are divergence-free, where

$$\vec{\nabla} \cdot \vec{F} \equiv 0$$

then we can readily see that there will be neither production nor destruction of field lines within this volume. Thus any internal field line will be conserved.

In such regions consider the field between two arbitrary surfaces A and B. (These would become curves or points for two- and one-dimensional fields, respectively).

Let us draw an arbitrary closed contour, a, on surface A. Through each point, α , on a, a field line will pass which will intersect the surface B at point β. As the point α is moved around curve a, the point will trace a corresponding closed curve, b, on B. Moreover, the field line segments, $\alpha\beta$, serve as generators or directrices of a tube T, which we shall define as a field tube.



In electromagnetism such tubes were first named by Michael FARADAY "sphondyloids" but James Clerk MAXWELL used the term "solenoid" from the Greek σωλη′ν (solen) "a tube". The current usage in electrodynamics is tube of induction or flux tube. In fluid mechanics such tubes are called stream tubes.

Everywhere in such tubes the field is conserved, resulting in a continuity relation. Moreover, by definition no FIELD VECTOR $\vec{F}$ can intersect a field tube.

The Third Derived Quantity

(The ROTATION of a VECTOR FIELD)

This derived operation → $\text{rot}\ \vec{F} = \vec{\nabla} \times F$ → Generates a VECTOR FIELD

The CURL or ROTATION is a DERIVED VECTOR QUANTITY equal to the LOCAL CIRCULATION per unit area.

We may compute the vector $\vec{G} = \vec{\nabla} \times \vec{F}$ very simply by matrix means; in Cartesian coordinates this outer product becomes:

$$\mathbb{G} = \nabla \times \cdot \mathbb{F}$$

$$\begin{bmatrix} G_1 \\ G_2 \\ G_3 \end{bmatrix} = \begin{bmatrix} 0 & -\partial/\partial z & +\partial/\partial y \\ +\partial/\partial z & 0 & -\partial/\partial x \\ -\partial/\partial y & +\partial/\partial x & 0 \end{bmatrix} \cdot \begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix}$$

The Stokes Circulation Theorem

Let us consider any closed contour C bounding an area A in any vector field $\vec{F}$.

Then the following fundamental theorem, due originally to Georg Gabriel STOKES, will always hold:

$$\boxed{\underset{\text{Circulation}}{\vec{\Gamma}} \equiv \underbrace{\oint_c \vec{F} \cdot \vec{dR}}_{\text{Boundary Evaluation}} \equiv \underbrace{\iint_A \vec{\nabla} \times \vec{F} \cdot \vec{dA}}_{\text{Surface Evaluation}}}$$

Irrotational Fields

If we know that $\vec{\Gamma}$ is identically Zero in some region, then we may always write

$$\vec{F} \equiv \pm \text{ grad } \phi = \pm \vec{\nabla}\phi$$

since $\vec{\nabla} \times \vec{\nabla}\phi \equiv 0$ for any scalar field. The ϕ so associated with $\vec{F}$ is called a scalar potential function, following the original terminology of George GREEN. The corresponding field is known as a potential field.

Conversely all gradient fields are irrotational by identity.

Solenoidal Fields

If we know that $\vec{G} \equiv \pm \vec{\nabla} \times \vec{F}$ then necessarily:

$$\vec{\nabla} \cdot \vec{G} \equiv 0$$

since $\vec{\nabla} \cdot \vec{\nabla} \times \vec{F}$ vanishes identically for any vector field $\vec{F}$. Such divergence-free $\vec{G}$ fields are called solenoidal.

Conversely, the vector $\vec{F}$ associated with $\vec{G}$ is said to be a vector potential function and generates a corresponding vector potential field.

calar and Vector Potential Fields

We may now summarize the results above for the dual cases of calar and vector potential functions.

IRROTATIONAL FIELDS
(Conservative)

If: $\vec{\nabla} \times \vec{f} = 0$

Then: $\vec{f} = \pm \overrightarrow{\text{grad}}\, e$

Since: $\vec{\nabla} \times \vec{\nabla} e \equiv 0$

hen e(s,y,z) is said to be a CALAR POTENTIAL FUNCTION which enerates a (SCALAR) POTENTIAL IELD.

aking $\vec{f} = + \overrightarrow{\text{grad}}\, e$ we have for he components

$$\left[\begin{array}{c} \partial e/\partial x \\ \hline \partial e/\partial y \\ \hline \partial e/\partial z \end{array}\right]$$

SOLENOIDAL FIELDS
(Source-free)

If: $\vec{\nabla} \cdot \vec{f} = 0$

Then: $\vec{f} = \pm \overrightarrow{\text{rot}}\, \vec{F}$

Since: $\vec{\nabla} \cdot \vec{\nabla} \times \vec{F} = 0$

Then F(x,y,z) is said to be a VECTOR POTENTIAL FUNCTION which generates a VECTOR POTENTIAL FIELD.

Taking $\vec{f} = + \overrightarrow{\text{rot}}\, F$ we have for the components

$$\left[\begin{array}{c} (\partial F_z/\partial y) - (\partial F_y/\partial z) \\ \hline (\partial F_x/\partial z) - (\partial F_z/\partial x) \\ \hline (\partial F_y/\partial x) - (\partial F_x/\partial y) \end{array}\right]$$

or the two-dimensional or planar field we find $F_x \equiv F_y \equiv 0$ and F_z may e taken as the scalar, f. Then we obtain the results:

$$\vec{f} = [\partial e/\partial x \vdots \partial e/\partial y] = [\partial f/\partial y \vdots -\partial f/\partial x]$$

This conjugate plane potential field is of singular importance o the analysis of all physical systems. Its relation to conformal apping we next discuss.

Plane Potentials and Conformal Maps

Many plane potential field problems may be solved readily employing functions of a complex variable. We outline here only a brief account of the resultant conformal mappings.

If the argument of any algebraic or transcendental function is a complex number $z = x + jy$, then, in general the function $w = f(z)$ will also be a complex number $w = u + jv$. If this function is analytic then

$$(\partial u/\partial x) + j(\partial v/\partial x) = f'(z)$$

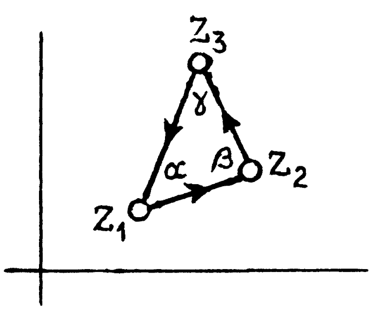
$$(\partial u/\partial y) + j(\partial v/\partial y) = jf'(x)$$

Comparing terms gives the CAUCHY-RIEMANN equations:

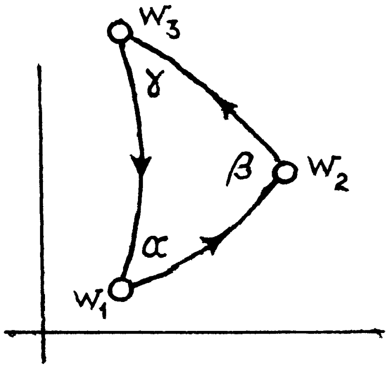
$$\partial u/\partial x = \partial v/\partial y \; ; \quad \partial u/\partial y = - \partial v/\partial x$$

But these conditions have two interesting consequences. First they ensure that every transformation $w = f(z)$ which satisfies these conditions will map portions of the z-plane into the w-plane such that relative directions and angles are preserved; such transformations are called conformal mappings and would map any set of three z-points as follows:

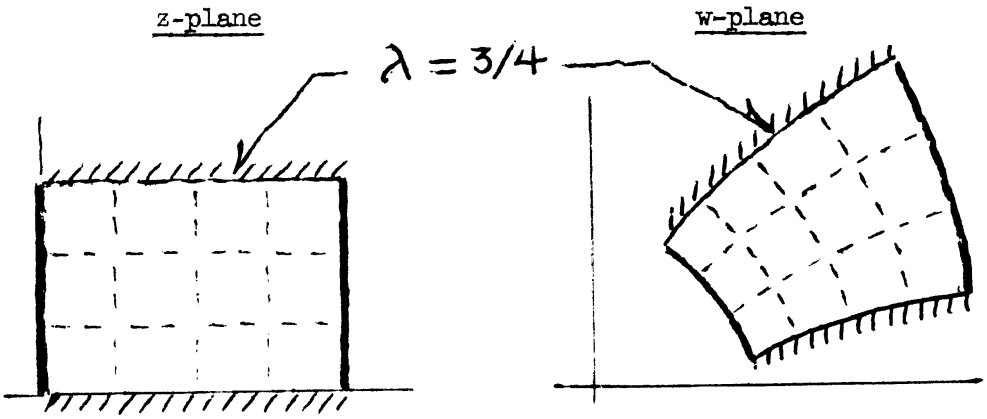
z-plane



w-plane



As a result of conformality a rectangular grid in the z-plane is mapped into an orthonormal meshwork as indicated.



The second interpretation of the CAUCHY-RIEMANN equations results from the fact that they are precisely equivalent to the results we obtained above for plane potential fields, namely

$$\vec{f} = \overrightarrow{\mathrm{grad}}\ e = [\partial e/\partial x \mid \partial e/\partial y]$$

also $$\vec{f} = \overrightarrow{\mathrm{rot}}\ F = [\partial f/\partial y \mid -\partial f/\partial x]$$

Thus any conformal mapping produces a plane potential field. We may always interpret one set of lines as equipotentials and the conjugate set as the corresponding fieldlines. However, and perhaps more importantly, it becomes apparent that the field form factor λ is a conformal invariant and represents nothing but the appropriate aspect ratio of the conformally equivalent rectangle! This fact we treat in more detail below.

Background Reading -- Conformal Mapping

(1) BECKENBACH, E. F., Editor: Construction and Application of Conformal Maps, Proceedings of a Symposium, National Bureau of Standards Applied Mathematics Series, Vol. 18 (1952).

Perhaps the single most useful reference in this subject.

(2) ROTHE, R; OLLENDORFF, F. and POHLEHAUSEN, K.: Theory of Functions, (1933)

A translation of a classical German text which predated American engineering application.

Background Reading -- *Conformal Mapping* (Continued)

(3) BEWLEY, L. V.: *Two-Dimensional Fields in Electrical Engineering* (194

An excellent treatment of the practical use of conformal maps to several fields of engineering.

(4) KOBER, H: *Dictionary of Conformal Representation* (1952)

A useful table in applying mapping functions.

(5) NEHARI, Zeev: *Conformal Mapping* (1952)

(6) CARATHEODORY, C.: Conformal Representation, Number 28, *Cambridge Tracts in Mathematics and Mathematical Physics* (1952)

The two above monographs deal with the analytical and formal properties of conformal maps, including 3-dimensional transformation.

D. Field Form Factors

As indicated earlier in this section, the transmittance of a homogeneous energetic field depends only upon the property constant of the medium and upon the field geometry in terms of (size) times (shape). This latter shape constant is measured by the *field form factor*, λ , which then can depend only upon the (normalized) boundary conditions. For a field tube, these boundary constraints consist of *fieldlines* along the walls and two terminal *isotimic* surfaces.

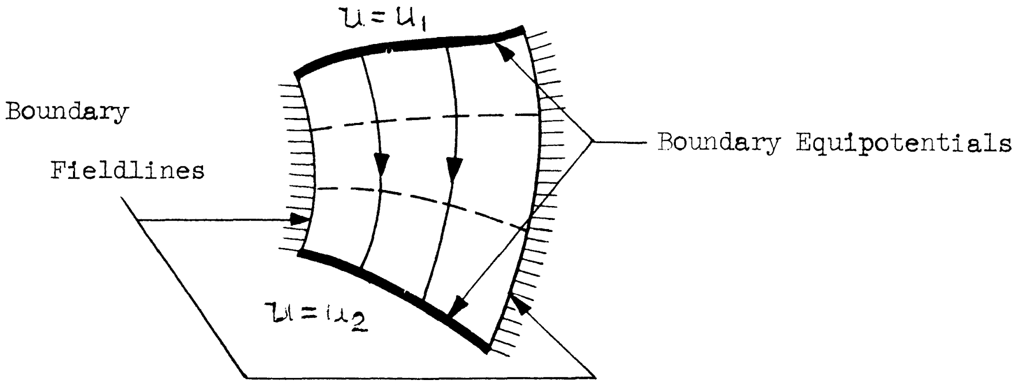
If the field is a *potential field*, the form factor is unique and invariant under conformal mappings. These properties of λ yield a number of simple bounds and estimating techniques as indicated below.

The Evaluation of Field Transmittances

The general field relations above may now be applied specifically to the problem of evaluation overall transmittances of a field tube. We assume in each case the existence of the following primitive vectorial relations:

A SCALAR POTENTIAL: $u = u(\vec{R}, t)$
The POTENTIAL GRADIENT: $\vec{U} = - \overrightarrow{\text{grad}}\, u$
An ASSOCIATED VECTOR: $\vec{V} = \kappa \vec{U}$

We may represent an arbitrary field tube in the following planar fashion:



Then along the tube between the two bounding equipotentials we have the relation:

$$\text{POTENTIAL DIFFERENCE:} \quad u_1 - u_2 = \int_1^2 \vec{U} \cdot d\vec{R}$$

while across the tube between the bounding fieldlines we have the relation:

$$\text{TOTAL FLUX:} \quad v = \int_A \vec{V} \cdot d\vec{A}$$

We are now in position to define the overall field transmittance in the fashion:

$$\text{FIELD TRANSMITTANCE:} \quad \mathcal{T} \equiv \frac{\int_A \vec{V} \cdot d\vec{A}}{\int_1^2 \vec{U} \cdot d\vec{R}} \equiv \frac{v}{u_1 - u_2}$$

In every field of physical origin, it is this quantity which is of greatest significance and importance at the macroscopic level. Historically, these phenomenological relations were discovered independently and generally are named after their discoverers (e.g. OHM's Law).

However, it is our present purpose to demonstrate that given the existence of an underlying relatively homogeneous and isotropic field, the corresponding transmittance may be evaluated from the field geometry and material transmissivity.

It may be readily demonstrated from dimensional reasoning that, since

$$\vec{U} = -\overrightarrow{\text{grad}}\, u \qquad \text{and} \quad \vec{V} = \kappa \vec{U} \text{ , then:}$$

$$\mathcal{T} = \kappa \frac{\int \vec{U} \cdot \vec{dA}}{\int \vec{U} \cdot \vec{dR}} = \kappa \frac{A_v}{L_u}$$

where A_v and L_u are an appropriate area and length, respectively, to absorb all information about the field geometry. But we may further factor the geometrical aspect into a (size) x (shape) product since

$$A_v/L_u = (A_n/L_n)\ \lambda' = (L^2/L)\ \lambda = L\ \lambda$$

Here A_n and L_n are arbitrary or nominal dimensions but L measures absolute size and λ measures the characteristic form.

Thus we obtain finally:

$$\boxed{\text{FIELD TRANSMITTANCE: } \mathcal{T} = \kappa \cdot L \cdot \lambda = \underbrace{(\text{Property})}_{\substack{\text{MATERIAL}\\ \text{Parameter}}} \cdot \underbrace{(\text{Size}) \cdot (\text{Form})}_{\substack{\text{GEOMETRICAL}\\ \text{Parameters}}}}$$

The value of λ depends only upon the form of the field--the boundary configurations and conditions--and not in any way upon the absolute scale or the particular medium. This fact is not only the principal justification for the practical use of field analogies; it also permits rapid estimates of transmittance by simple inspection of field geometry.

For two-dimensional or planar fields the dimension L can be taken as the transverse width of the medium, giving a particularly simple result for the transmittance per unit width, namely:

$$\boxed{\text{PLANAR SPECIFIC TRANSMITTANCE: } \mathcal{T}/L = \kappa \cdot \lambda}$$

This relation clearly implies that geometrically similar planar fields have identical form factors. But the same result is also true for all potential fields.

Energy Storage and Dissipation in a Potential Field

Given the steady potential field e(x,y,z) and the associated vector field $\vec{f}(x,y,z) = -k\,\overrightarrow{\text{grad}}\,e$, the field energy contained in an arbitrary closed region is given by:

$$\mathbb{E} = \int_{\mathcal{V}} |\vec{f}|^2 \, d\mathcal{V} = k \int_{\mathcal{V}} |\overrightarrow{\text{grad}}\, e|^2 \, d\mathcal{V}$$

But using the divergence theorem these volume integrals may be directly equated to the surgace integral to obtain

$$\mathbb{E} = \int_A e\vec{f}\cdot d\vec{A} = k \int_A e\, \overrightarrow{\text{grad}}\, e \cdot d\vec{A}$$

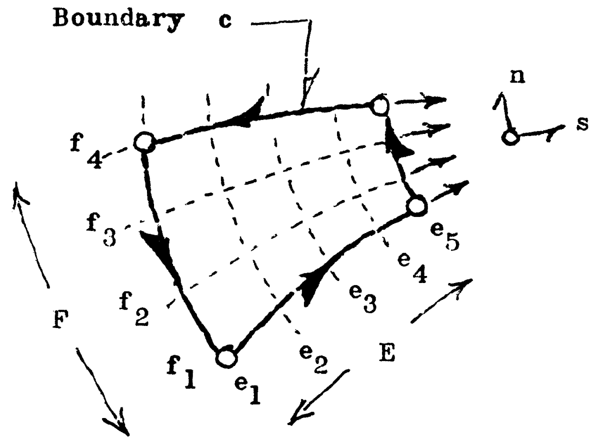
In the discussion below, we shall restrict attention to the special case of plane potential fields. However the methods and relations derived are of general validity.

For this case, the above general principles reduce to the following relations:

$$\mathbb{E}/L = k \oiint_A \left[\left(\frac{\partial e}{\partial x}\right)^2 + \left(\frac{\partial e}{\partial y}\right)^2\right] \cdot dxdy = k \oiint_A \left[\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2\right] dxdy$$

$$= k \oint_c e \frac{\partial e}{\partial n} \, ds = k \oint_c e \frac{\partial f}{\partial s} \, ds = k \oint_c e \, df$$

The form expressed at the lower right is most easily evaluated alone e-lines and f-lines as indicated below:



Thus we obtain for the energy per unit width:

$$\mathbb{E}/L = k \oint e\,df = k\,(e_5 - e_1)\cdot(f_4 - f_1)$$

$$= k \cdot E \cdot F$$

It is obvious now that this is merely the simple summation of the equal energies distributed in each orthonormal cell. Moreover, since $F = \lambda\, E$ then we obtain finally

$$\boxed{\mathbb{E} = k \cdot L \cdot \lambda \cdot E^2 = k \circ L \circ \frac{1}{\lambda} \cdot F^2}$$

Thus the form factor λ not only serves as a measure of the field transmittance; it also is a direct measure of the energy storage or dissipation in any potential field.

Field Energy Extremum Principles

Returning to the volumetric evaluation of field energies we are now in a position to state a pair of complementary extremum principles which may be used to determine the field energy and therefore the form factor. These we shall name after their promulgators, namely:

DIRICHLET's Principle

Of all arbitrary scalar fields in a region satisfying a given boundary condition, the POTENTIAL field has the MINUMUM energy.

THOMSON's Principle

Of all solenoidal vector fields in a region satisfying a given boundary condition, the IRROTATIONAL field has the MINIMUM energy.

The two principles may then be used very effectively to estimate upper and lower bounds for the field form factor to obtain results in the form:

$$\lambda_D \geqq \lambda \geqq \lambda_T$$

For DIRICHLET's principle we make an arbitrary assignment of isotimics compatible with the boundary conditions; for THOMSON's principle, the fieldlines are assumed. If, and only if, the assumptions correspond to equipotentials and gradient lines, respectively, will the corresponding field energies be minimized.

For some of the interesting history behind these dual principles we may quote POLYA from the work cited below:

> An important special case of Thomson's principle... was already known to Gauss...That the two principles can be used to obtain estimates of the capacity from opposite sides, has been observed by Maxwell to whom the method was suggested by an investigation of Lord Rayleigh...Maxwell, however, did not derive upper or lower bounds from his method in concrete cases, observing that the "operations...that are in general too difficult for practical purposes".

That this is not quite the case, particularly in this era of machine computation, is amply testified to in the book of POLYA and also in CRANDALL's text which is referenced below:

It is perhaps easiest to visualize this application for the case of dissipation in a planar field. However analogous developments are possible for all energetic fields.

DIRICHLET Principle	THOMSON Principle
Let $e(x,y)$ be an arbitrary scalar function such that $e \equiv 1$ on surface A of a fieldtube and $e \equiv 0$ on surface B. Then $E \equiv 1$.	Let $f(x,y)$ be an arbitrary source-free vector function such that the fieldlines are normal to A and B, coincide with the fieldtube boundaries, and enclose one unit of flow. Then $F \equiv 1$.

$$\mathbb{P}_D \equiv \rho \int_{\mathcal{V}} |\overrightarrow{\text{grad}}\, e|^2 \, d\mathcal{V} \quad \Big\vert \quad \mathbb{P}_T \equiv \rho \int_{\mathcal{V}} |\vec{f}\,|^2 \, d\mathcal{V}$$

$$\equiv G_D \cdot E^2 = G_D \quad \Big\vert \quad \equiv R_T \cdot F^2 = R_T$$

But the above minimum principles tell us

$$G_D \geq G \quad \Big\vert \quad R_T \geq R$$

and the equality signs hold only for:

$$\nabla^2 e \equiv 0 \quad \Big\vert \quad \vec{\nabla} \times \vec{f} \equiv 0$$

Thus we may obtain upper and lower bounds in the form:

$$\boxed{\begin{array}{c} G_D \geq G \geq G_T \\ R_D \leq R \leq R_T \end{array}}$$

If the bounds are close then

$$\boxed{\begin{array}{c} G \cong \sqrt{G_D \cdot G_T} \\ R \cong \sqrt{R_D \cdot R_T} \end{array}}$$

(Note in the above approximation if arithmetic means were used different estimates would be obtained for conductance and resistance, which is not particularly logical.)

Background Reading -- Dirichlet - Thomson Principles

(1) POLYA, G. and SZEGO, G.: Isoperimetric Inequalities in Mathematical Physics. Annals of Mathematics Studies, Number 27 (1951).

(2) COURANT, R.: Dirichlet's Principle, Conformal Mapping, and Minimal Surfaces. (1950).

Maxwell's Principle of Transmissivity

A universal field principle of great utility was first ennunciated by James Clerk MAXWELL. It may be stated in general dualistic terms in the following words.

In any field, if the transmissivity is locally

DECREASED		INCREASED
the overall	transmittance	is correspondingly
DECREASED		INCREASED

It then follows that fixing any

fieldline	equipotential
DECREASES	INCREASES

the overall field transmittance.

These last conditions amount to a loosening and generalization of the

THOMSON Principle	DIRICHLET Principle

respectively.

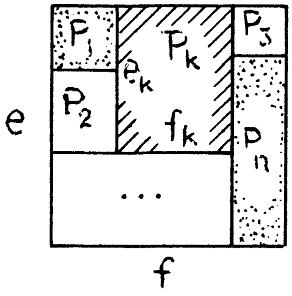
The transmissivity principle has practical application to uniform and nonuniform, linear and nonlinear, continuous and reticular fields of all sorts. As indicated above, it may provide bounded estimates of the form factors and transmittances of fields; but, in addition, by judicious compensatory or balanced increases and decreases of transmissivity (or local transmittance) throughout the extent of a field a rational estimate may be readily obtained.

E. Rectangle Diagrams

In the earlier discussion of state-determined elements, the concept of complementary energies was introduced. While these notions originated with MAXWELL and HELMHOLTZ in the last century, only recently a striking interpretation was given to the distribution of generalized energy over reticular fields in terms of rectangle diagrams. These figures, introduced chiefly through the efforts of CHERRY and MILLAR, represent the generalization of the concept of form factor to arbitrary linear and nonlinear fields. They serve to portray graphically the local distribution of stored or dissipated energy over the extent of a reticular 1-port system or network of like-kind 1-port elements (i.e. all-resistor, all-capacitor, or all-inertor reticular fields).

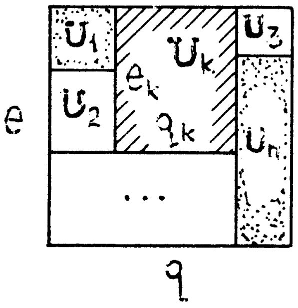
For each of the three species of nets, the rectangle diagrams are mosaics as follows:

RESISTIVE FIELD



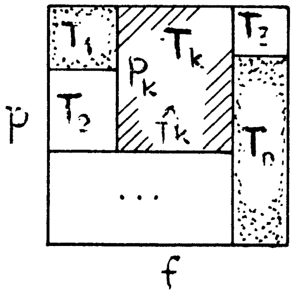
$$\mathbb{P} \equiv \sum_k P_k$$

CAPACITIVE FIELD



$$\mathbb{U} \equiv \sum_k U_k$$

INERTIAL FIELD



$$\mathbb{T} \equiv \sum_k T_k$$

Since the appropriate generalized energies are all compatibly distributed, so must also be the corresponding normal and complementary energies. For a uniform continuous field, with unit property constant, the rectangle diagram simply becomes the conformal mapping of the field into an equivalent rectangle of identical form factor; as before, the form factor, λ , measures the aspect ratio of this rectangle, all of whose elements are equal squares.

However, the importance of the rectangle diagram is rather for depicting the case of nonuniform and nonlinear reticular fields, where the elements themselves will also generally all be diverse but compatible and contingent rectangles. Clearly, the micro-reticulation can then extend indefinitely to the field substructure if desirable, but this is not a necessary requirement.

Dualistic Energy Principles for Static and Stationary Fields

Whenever we deal with continuous or reticular, linear or nonlinear fields of like-kind elements operating under steady conditions, the dual pair of Lagrange Equations cited previously yield a simple set of dual extremum principles for each type of element as follows:

CAPACITANCE:	$\partial \mathbb{U}_q / \partial \mathbb{q} \equiv 0$	$\partial \mathbb{U}_e / \partial \mathbb{e} \equiv 0$
RESISTANCE:	$\partial \mathbb{P}_f / \partial \mathbb{f} \equiv 0$	$\partial \mathbb{P}_e / \partial \mathbb{e} \equiv 0$
INERTANCE:	$\partial \mathbb{T}_f / \partial \mathbb{f} \equiv 0$	$\partial \mathbb{T}_p / \partial \mathbb{p} \equiv 0$

Since all these pairs have identical topological structure, we shall center attention on the resistance case. These dual extremum conditions give us simple generalizations to arbitrary fields of the classical

THOMSON Principle	DIRICHLET Principle

discussed above.

For the resistance case, their significance is the following: in any reticular field embedded in a fixed environment, where all bounding parts have been represented as equivalent:

Neumann (Helmholtz) Ports	Dirichlet (Thevenin) Ports

for any arbitrary internal assignment of

Flow $\mathbb{f}$	Effort $\mathbb{e}$

compatible with the boundary conditions, that set which minimizes total

Content $\mathbb{P}_f$ Cocontent $\mathbb{P}_e$

represents the equilibrium or steady-state configuration. Similar statements hold for capacitive and inertial fields.

Not only do these principles offer a particularly attractive way to estimate the steady-state solution for such fields but they also lead directl as before to bounded estimates of overall transmittance.

Again, for the resistance case an approach to the minimum for:

CONTENT COCONTENT
will furnish an approximate
LOWER BOUND UPPER BOUND

for the overall conductance (≡ resistive transmittance).

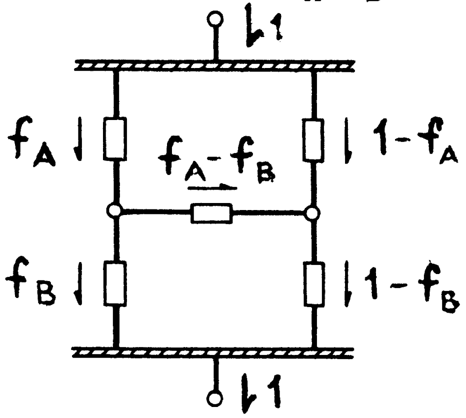
These properties may be made self-evident through use of rectangle diagrams as indicated below.

Energy Extrema in Terms of Rectangle Diagrams

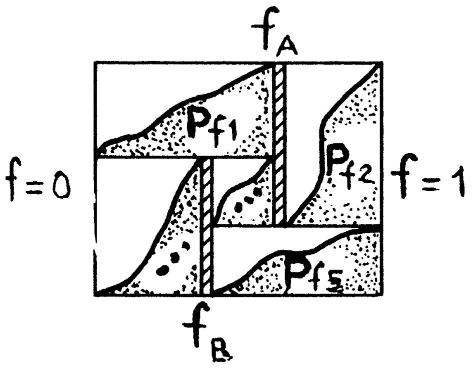
We may demonstrate the relation of the above principles to rectangl diagrams in terms of a loaded resistive Wheatstone bridge. For simplicity, l us consider the case where the variables have been normalized as indicated. broader application of these ideas should be obvious from this example.

GENERALIZED THOMSON PRINCIPLE

Assignment of (f_A, f_B)



Rectangle Diagram:



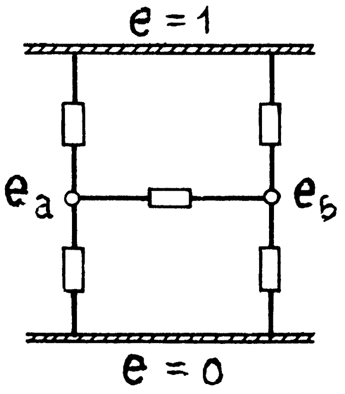
$$\mathbb{P}_f \geq \mathbb{P}_{f,\text{ equil.}}$$

For the linear case:

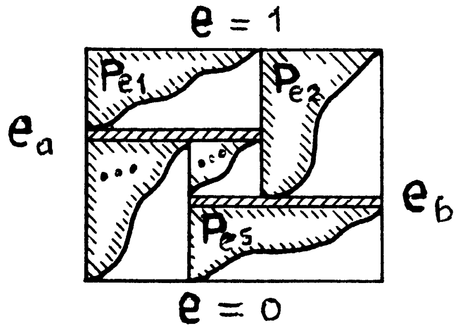
$$R_f \geq R$$

GENERALIZED DIRICHLET PRINCIPLE

Assignment of (e_a, e_b)



Rectangle Diagram:



$$\mathbb{P}_e \geq \mathbb{P}_{e,\text{ equil.}}$$

For the linear case:

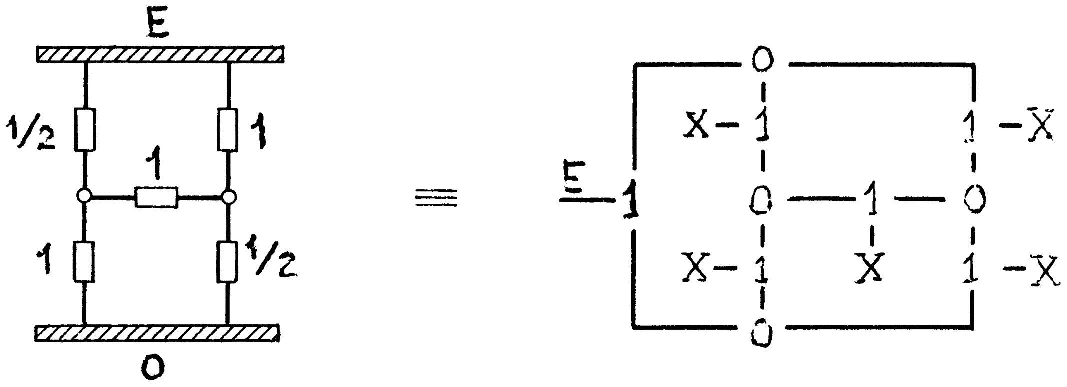
$$G_e \geq G$$

$$\boxed{\begin{array}{ccccc} R_f & \geq & R & \geq & R_e \\ G_f & \leq & G & \leq & G_e \end{array}}$$

In the above relations, the equality signs hold only for the equilibrium case, corresponding to a correct assignment of system variables.

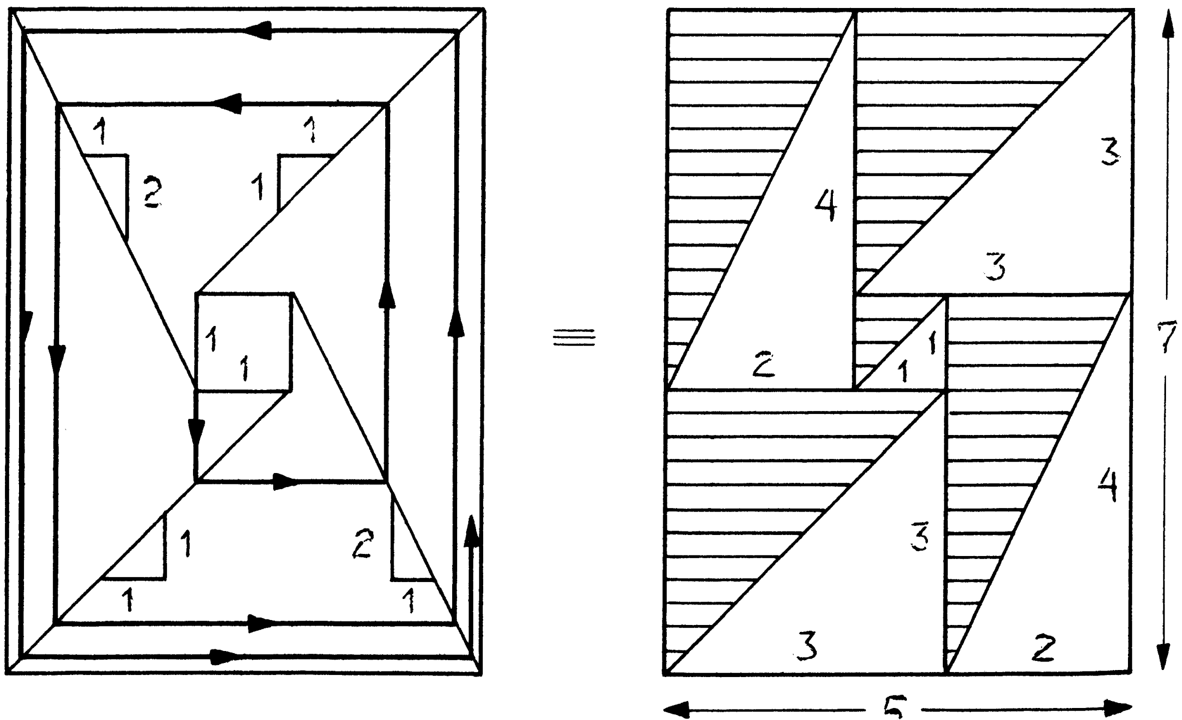
The Coons Construction

As another example of the application of rectangle diagrams, consid the problem of determining the overall transmittance of the following bridge structure:



where the local linear transmittances are as indicated. The solution of all such linear (and nonlinear !) bridges may be elegantly determined using a simple rectangle diagram construction originating with Stephen A COONS as follows:

Coon's Construction Rectangle Diagram



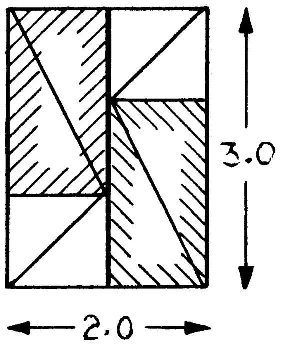
his gives a final transmittance:

$$\mathcal{T} = 5/7 = \underline{0.71}$$

This result may be check using the MAXWELL Transmissivity Principle s follows:

LOWER BOUND

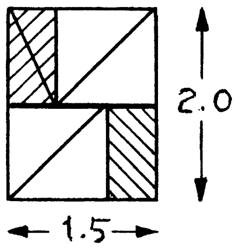
Zero Central Transmittance



$$\mathcal{T} \geqq 2/3 = 0{\cdot}67$$

UPPER BOUND

Infinite Central Transmittance



$$\mathcal{T} \geqq 1{\cdot}5/2{\cdot}0 = 0{\cdot}75$$

$$\boxed{\mathcal{T} \cong \sqrt{\frac{2}{3}\cdot\frac{3}{4}} = 0{\cdot}707}$$

ackground Reading -- Rectangle Diagrams

1) CHERRY, E. C. and MILLAR, W.: Some New Concepts and Theorems Concerning Non-linear Systems, Automatic and Manual Control, pp. 263-274 (1952).

2) GARDNER, Martin: Mathematical Games, Scientific American, Vol. 199, Number 5, pp. 136-142 (November, 1958)

Background Reading -- Potential Functions

(1) GREEN, George: An Essay on the Application of Mathematical Analysis to the Theories of Electricity and Magnetism (1828).

(2) -------------: Mathematical Investigations concerning the Laws of Equilibrium of Fluids analogous to the Electric Fluid, with other similar Researches. Phil. Trans. (1833).

These two papers by a self-taught genius (and miller!) are the foundations of modern potential theory.

(3) KELLOGG, O. D.: Foundations of Potential Theory (1929).

(4) MACMILLAN, W. D.: The Theory of the Potential, Second Edition (1958).

The above two books are the classic references.

Background Reading -- Fields

(1) MAXWELL, J. C.: A Treatise on Electricity and Magnetism, Third Revised Edition, Vol. I and II (1891).

Maxwell's contribution to field theory ranks with Newton's gift to dynamics.

(2) MASON, Max and WEAVER, Warren: The Electro-Magnetic Field (1929)

A definitive work of an analytical nature.

(3) SKILLING, H. H.: Fundamentals of Electric Waves (1942)

An outstanding lucid introduction to field phenomena treated by vectorial mechanics.

(4) WEBER, Ernst: Electromagnetic Fields, Vol. I - Mapping of Fields Second Edition (1954).

Very instructive and readable introduction to field theory.

(5) ROGERS, W. E.: Introduction to Electric Fields (1954).

Especially elegant graphical and analog field plots.

(6) SCHNEIDER, P. J.: Conduction Heat Transfer (1955)

An excellent treatment of transient and steady field phenomena disguised under the rubric of an engineering speciality.

(7) BITTER, Francis: Currents, Fields and Particles (1956)

A text which has been used in the basic undergraduate curriculum at M.I.T.

F. The Temporal Response of Physical Systems

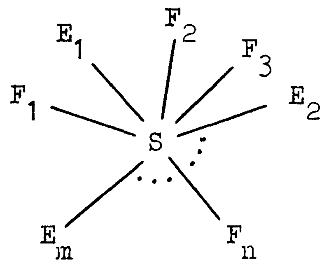
Let us use the closed, state-determined system as an example. The conditions at the boundary between system and environment can always be at least approximately represented in terms of set of (possibly mixed) Thevenin and Helmholtz equivalent sources, where:

Thevenin Source	Helmholtz Source
$E - \underset{\underset{Z}{\bullet}}{1} -$	$F - \underset{\underset{Z}{\bullet}}{0} -$

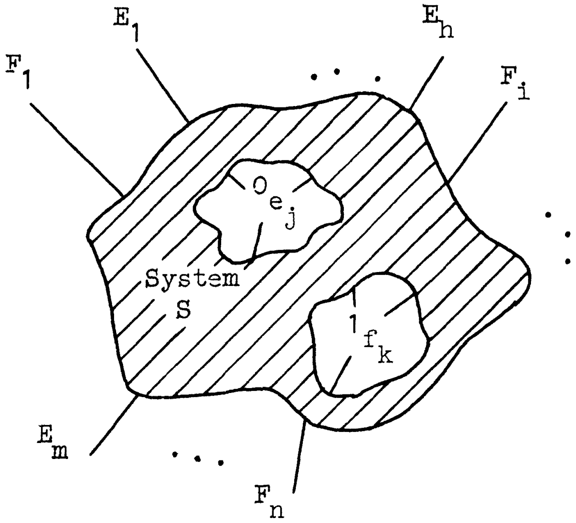
We may then further idealize these sources by extending the <u>prime reticulation</u> -- the fixing of the system boundary -- to include the source impedances as parts of the system itself in the fashion:

For both continuous and reticular fields this strategem results in the classical boundary conditions of Dirichlet or of Neumann, or else mixtures of these forms. The specifications of $E = E(t)$ and $F = F(t)$ of course include constants and null values as important special cases.

Let us now consider a multiport state-determined system S which has been rendered into such a form:



In general each <u>internal</u> state variable will be determined adjacent to an appropriate energy junction; this situation we might indicate in the fashion:



On the above diagram the (m + n) ideal boundary sources are all independent variables or inputs while the internal energetic variables e_j and f_k are each and every one dependent or output variables. We may then indicate the final explicit functional solutions in the form:

$$\begin{aligned} &\vdots && \vdots \\ e_j(t) &= \Psi_j\,[\,E_1(t), \ldots, E_m(t);\ F_1(t), \ldots, F_n(t)\,] \\ &\vdots && \vdots \\ f_k(t) &= \Psi_k\,[\,E_1(t), \ldots, E_m(t);\ F_1(t), \ldots, F_n(t)\,] \\ &\vdots && \vdots \end{aligned}$$

We are merely stating that the instantaneous value of any internal state variable is a function only and entirely of the history of the ideal environment. We should note that the output variables conjugate to each external input source are now determined from within the system and are therefore included in the set $[e_j\ ,\ f_k]$.

It is also interesting to note the singular result which arises for the special case of a constant environment. In this case, for a state-determined system, the response trajectory is uniquely determined by the initial state, alone.

The Response of Linear Systems

A system is linear if all the elemental components of a system are linear, in the sense that the governing relationships among the energy variables are all linear. In the case of the state determined systems, the primitive characteristics are linear static functions while for the more general n-port linear elements, the reticulation may be carried only to the level of linear functional operators.

If the functionals Ψ_j and Ψ_k are all linear operators then the superposition property implies the additional functional reticulation:

$$\begin{array}{ccccccccc} \vdots & & & \vdots & & & \vdots & & \\ e_j & = & \cdots & + \mathbb{F}_{jh} E_h & + & \cdots & + \mathbb{Z}_{ji} F_i & + & \cdots \\ \vdots & & & \vdots & & & \vdots & & \\ f_k & = & \cdots & + \mathbb{Y}_{kh} E_h & + & \cdots & + \mathbb{F}_{ki} F_i & + & \cdots \\ \vdots & & & \vdots & & & \vdots & & \end{array}$$

The operators $\mathbb{F}_{jh}$ and $\mathbb{F}_{ki}$, being dimensionless, are termed transfer ratios, while the dimensioned operators $\mathbb{Z}_{ji}$ and $\mathbb{Y}_{kh}$ are called transfer impedances and transfer admittances, respectively; all four classes of operators may be considered transfer functions.

Using scaling constants, the transfer impedances and admittances may also be expressed in pure ratio form; for example:

$$\mathbb{Z}_{ji} = R\,\mathbb{F}_{ji} = (E_o/F_o)\,\mathbb{F}_{ji}$$
$$\mathbb{Y}_{kh} = G\,\mathbb{F}_{kh} = (F_o/E_o)\,\mathbb{F}_{kh}$$

By this means all the transfer operators reduce to dimensionless operators $\mathbb{F}_{ab}$, or to such ratios multiplied or divided by a nominal resistance constant. We may then restrict attention to the single element:

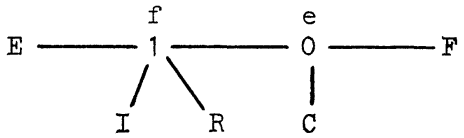
$$y = \mathbb{F}\,x \qquad \text{or} \qquad \xrightarrow{x}\boxed{\mathbb{F}}\xrightarrow{y}$$

Transfer Characteristics of Linear State-Determined Reticular Systems

Such systems are reticular complexes of the following seven linear or linearized elements:

$$[\ E\cdot,\ F\cdot,\ R\cdot,\ C\cdot,\ I\cdot,\ \cdot\dot{0}\cdot,\ \cdot\dot{1}\cdot\ ]$$

A simple example involving all seven might be the system:



The general functional relations for e and f would be written:

$$f = \Psi_1 (E, F)$$
$$e = \Psi_2 (E, F)$$

which in turn linearize to the form:

$$f = (1/R) \cdot \mathbb{F}_{11} \cdot E + \mathbb{F}_{12} \cdot F$$

$$e = \mathbb{F}_{21} \cdot E + R \cdot \mathbb{F}_{22} \cdot F$$

using the resistance R as the scaling constant.

Any of a wide variety of linear reduction schemes will yield for the above system:

$$\mathbb{F}_{11} = \frac{(RC)D}{1 + (RC)D + (IC)D^2} \; ; \; \mathbb{F}_{12} = \frac{1}{1 + (RC)D + (IC)D^2}$$

$$\mathbb{F}_{21} = \frac{1}{1 + (RC)D + (IC)D^2} \; ; \; \mathbb{F}_{22} = - \frac{(I/R)D + 1}{1 + (RC)D + (IC)D^2}$$

Since each operator is dimensionless, and, indeed, the numerator and denominators are separately nondimensional, the coefficients of the powers of D are accordingly time constants raised to the corresponding power.

Thus the three physical parameters (R,C,I) manifest their effects through the three time constants (T_1, T_2, T_3) where

$$T_1 \equiv RC \quad ; \quad T_2 \equiv \sqrt{IC} \quad ; \quad T_3 \equiv I/R$$

In these terms the above operators become:

$$\mathbb{F}_{11} = \frac{T_1D}{1 + T_1D + T_2^2D^2} \quad ; \quad \mathbb{F}_{12} = \frac{1}{1 + T_1D + T_2^2D^2}$$

$$\mathbb{F}_{21} = \frac{1}{1 + T_1D + T_2^2D^2} \quad ; \quad \mathbb{F}_{22} = \frac{-(T_3D + 1)}{1 + T_1D + T_2^2D}$$

The fact that $\mathbb{F}_{21} \equiv \mathbb{F}_{21}$ is a consequence of the reciprocity principle holding for all such linear passive systems.

From the above we may see that the general case of a linear system having r external bonds to the environment and a total of s additional state variables, will entail (r + s) energy junctions whose (r + s) output states $\mathbb{Y}$ may be related to the r input variables $\mathbb{X}$ by the linear matrix form:

$$\mathbb{Y} = \mathbb{T} \cdot \mathbb{X}$$

$$\begin{array}{c} r + s \\ \text{rows} \end{array} \left\{ \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_r \\ \vdots \\ y_{r+s} \end{bmatrix} \right. = \underbrace{\begin{bmatrix} \mathbb{F}_{11} & \mathbb{F}_{12} & \cdots & \mathbb{F}_{1r} \\ \mathbb{F}_{21} & \mathbb{F}_{22} & \cdots & \mathbb{F}_{2r} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbb{F}_{r1} & \mathbb{F}_{r2} & \cdots & \mathbb{F}_{rr} \\ \vdots & \vdots & & \vdots \\ \mathbb{F}_{r+s,1} & \mathbb{F}_{r+s,2} & \cdots & \mathbb{F}_{r+s,r} \end{bmatrix}}_{r\ -\ \text{columns}} \cdot \left. \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{bmatrix} \right\} \begin{array}{c} r \\ \text{rows} \end{array}$$

Background Reading -- Linear Systems

(1) CHENG, D. K.: Analysis of Linear Systems (1959)

(2) TRIMMER, J. D.: Response of Physical Systems, Second Printing (1953)

G. System Response in the Time and Frequency Domains

The Description of Arbitrary Signals

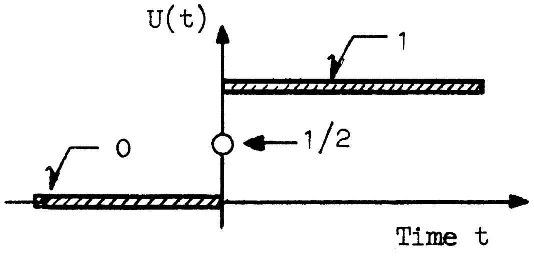
Under normal operating conditions the variables of most physical and engineering systems will undergo arbitrary variations over time. The general situation will involve stochastic signals, the word deriving from the Greek στωχαστικος (stochastikos) meaning, surprisingly enough, both "to aim" and "to guess". Such variable signals are those which have some probabilistic element and are thus not completely deterministic. At the extremes of the stochastic range, we find the purely deterministic (i.e. point-predictable) signals at the one end and purely random (i.e. distribution-predictable) signals at the other end.

Since any deterministic functional operator, Ψ, applied to a stochastic signal, $\mathbb{X}$, will produce another stochastic signal $\mathbb{Y}$, we are necessarily concerned in all systems with an adequate description of arbitrary stochastic signals.

The detailed description of random signals and processes we leave to be considered in 2.752; here we shall concern ourselves with purely deterministic but otherwise arbitrary signals.

Consider first the unit step or jump function, U(t):

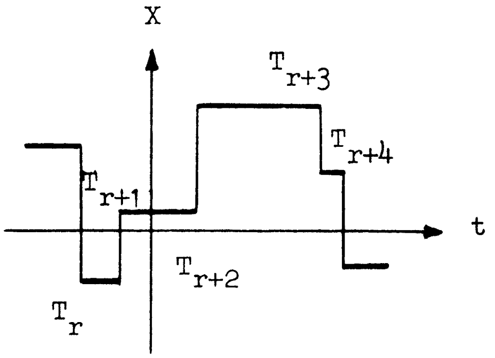
$$\boxed{U(t) \equiv \frac{1}{2}\ [1 + \text{sgn } t]}$$



This discontinuous but nevertheless analytic function was first introduced into system analysis by Oliver HEAVISIDE and is frequently called in his honor the Heaviside function.

A completely arbitrary stepwise varying signal can be defined by the equation:

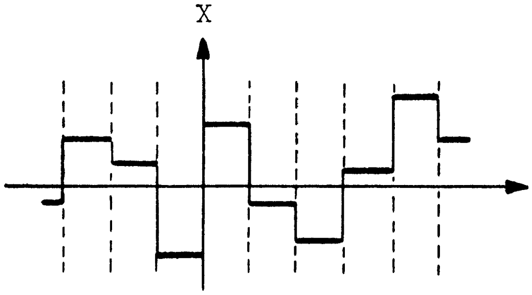
$$\boxed{X(t) \equiv \sum_{k=1}^{\infty} a_k U(t-T_k)}$$



Here the coefficients, a_k, and jump times, T_k, are assumed at will (subject to the implicit ordering; $T_k < T_{k+1}$).

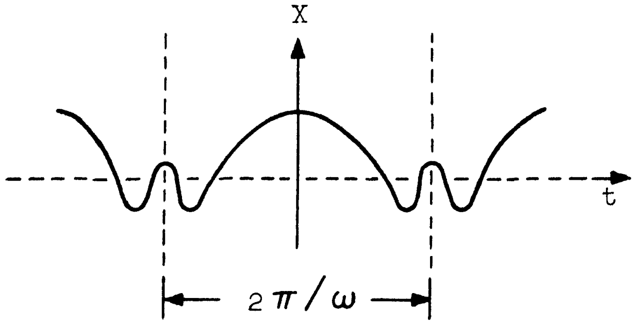
If we wish the jumps to be at synchronous clock intervals, T, we need merely to set $T_k \equiv kT$, to obtain:

$$\boxed{X(t) \equiv \sum_{k=-\infty}^{\infty} a_k U(t-kT)}$$



Now the sequence [a_k] is a precise specification of any such X(t), given a specified clock interval, T.

The significance of the above results lies in the fact that such descriptions in the time domain are the precise equivalents of the conventional Fourier series in the frequency domain. To demonstrate this let us for simplicity consider an <u>even</u> periodic function. This would be sketched and expanded as indicated



$$\boxed{X(t) \equiv \sum_{k=0}^{\infty} b_k \cos k\omega t}$$

Now by analogy to the last expansion above we have

Clock Interval: $T \longleftrightarrow \omega$: Fundamental Frequency
Jump Amplitude: $a_k \longleftrightarrow b_k$: Harmonic Amplitude
Jump Time : $kT \longleftrightarrow k\omega$: Harmonic Frequency

Of course the general expression for a Fourier series will involve both <u>odd</u> and <u>even</u> terms since <u>any</u> function can be expressed as the sum of an odd and an even function. Thus the general Fourier series is written in any of the equivalent forms:

$$X(t) = \sum_k a_k \sin (k\omega t + \theta_k)$$
$$= \sum_k (a_k \sin k\omega t + b_k \cos k\omega t)$$
$$= \sum_k b_k \cos (k\omega t + \phi_k)$$

The angles θ_k and ϕ_k are called the <u>harmonic phases</u>. We shall generally find the last form most convenient in use.

It is also not without interest to speculate on the frequency analog of the asynchronous arbitrary stepwise time signal above. This may be postulated as the (generally) aperiodic signal:

$$X(t) \equiv \sum_k b_k \cos (\omega_k t + \phi_k)$$

This is clearly a generalization of the Fourier series which will frequently give the appearance of a quasi-random signal if the set of $[\omega_k]$ are relatively incommensurable.

Thus with either the set $[a_k, T_k]$ or the set $[b_k, \omega_k, \phi_k]$ we may describe to an arbitrary degree of precision any normally encountered X(t) over a finite period of time. We are now in possession of a descriptive mechanics sufficient to treat the dynamics of any linear system either in the time or the frequency domain.

Superposition Properties of Linear Systems

Any functional operator, Ψ, is said to be a linear operator Λ, if it satisfies the linearity condition:

$$\boxed{\Lambda(\mathbb{a}\mathbb{X} + \mathbb{b}\mathbb{Y}) \equiv \mathbb{a}\Lambda\mathbb{X} + \mathbb{b}\Lambda\mathbb{Y}}$$

for arbitrary constant matrices $\mathbb{a}$ and $\mathbb{b}$. Of course the vector variables $\mathbb{X}$ and $\mathbb{Y}$ must be such as to lie within the domain of definition of Λ.

The above condition is both necessary and sufficient. Thus if the operator is "known" to be linear, then the linearity condition is necessarily satisfied. The above condition is customarily written for scalars and in the form of two simpler conditions as follows:

$$\boxed{\begin{array}{c} \Lambda(X_1 + X_2) \equiv \Lambda X_1 + \Lambda X_2 \\ \Lambda aX \equiv a\Lambda X \end{array}}$$

It is perhaps easier to see now that linearity implies only that Λ is distributive with respect to addition and commutative with respect to scalar multiplication.

For the systems of interest to us, we are concerned in the scalar case with situations.

$$Y = \Psi[X]$$

But from the above we know that:

$$\text{If:} \quad Y = \Lambda X$$

$$\text{and:} \quad X = \sum_k c_k X_k$$

$$\text{then:} \quad Y = \sum_k c_k \Lambda X_k$$

Thus *the response of a linear system to a disturbance composed of a weighted sum of signals is the correspondingly weighted sum of the responses to each signal acting alone*.

This superposition principle for linear systems is the basis for the application of nearly all mathematical and scientific theory to the real world. Up until its promulgation by Daniel BERNOULLI in connection with the vibrating string problem in 1753 the practical application of mathematics and analysis was severely restricted. The first dramatic results after its enunciation were the masterful trigonometric series expansions of Joseph FOURIER in 1822.

The superposition property is obviously also valid for linear multi-port systems in the form

$$\mathbb{Y} = \Lambda \mathbb{X}$$

$$\mathbb{X} = \sum_k \mathbb{C}_k \mathbb{X}_k$$

$$\mathbb{Y} = \sum_k \mathbb{C}_k \Lambda \mathbb{X}_k$$

In the paragraphs below we indicate the application of the superposition principle first to the representation of behavior in the time domain and, secondly, to representation in the frequency domain.

Step Response and Time Domain Behavior

Consider the time-invariant linear transfer operator $\mathbb{F}(D)$. We define the step response, $F(t)$, as

$$F(t) \equiv \mathbb{F}(D)\, U(t)$$

Thus the step response is merely the output Y(t) resulting from an input $X(t) \equiv U(t)$. We may then determine the response to an arbitrary jump function by using the superposition property as follows:

$$\text{Given:} \quad Y(t) = \mathbb{F}\, X(t)$$

$$\text{and:} \quad X(t) = \sum_k a_k\, U(t-T_k)$$

$$\text{and:} \quad F(t) = \mathbb{F}\, U(t)$$

$$\text{Then:} \quad F(t-T_k) = \mathbb{F}\, U(t-T_k) \qquad \text{[time invariance]}$$

$$\text{and:} \quad \mathbb{F}\, X(t) = \sum_k a_k\, \mathbb{F}\, U(t-T_k) \qquad \text{[superposition]}$$

$$\text{Therefore:} \quad \boxed{Y(t) = \sum_k a_k\, F(t-T_k)}$$

This result implies that for linear operators a knowledge of the step response alone is adequate to determine the behavior in the time domain to an arbitrary degree of precision. It is this fact which has raised the step response to the eminence which it has held for the last seventy years or more. Of course, if the system is essentially nonlinear and the normal input is arbitrary then the step response has little if any value as a behavioral measure; it is worth stressing this point in the light of contemporary proclivities for obtaining such meaningless data.

If we wish to pass to the limiting case of a smooth X(t) then the sums must go over into integrals. This transition is natural if Stieltjes integration is employed. Thus we define X in the form:

$$X(t) \equiv \int_{-\infty}^{t} dX(\tau)$$

If X is (purely) continuous in real time, t, (and ∴ umbral time, τ) then this reduces to the Riemanian identity:

$$X(t) = \int_{-\infty}^{t} dX(\tau) = X(t)$$

But if X is (purely) discontinuous at a series of discrete times [T_k] then the integral is evaluated as the sum of jumps or salti in X up to time t, namely:

$$X(t) = \sum_{k}^{T_k \leq t} a_k U(t-T_k)$$

Having established this relation we may then state the integral form of supe position as convolution integral:

$$\boxed{Y(t) = \int_{-\infty}^{t} F(t-\tau)\, dX(\tau)}$$

Due to linearity, it is simple to demonstrate that a complementary form of this convolution integral exists, namely:

$$\boxed{Y(t) = \int_{-\infty}^{t} X(t-\tau)\, dF(\tau)}$$

Furthermore, if $F(\tau)$ has no discontinuities (including one at the origin!) then we may usefully introduce the concept of the impulse response or weighting function, f(t), where

$$f(t) \equiv dF(t)/dt \quad \text{or} \quad F(t) \equiv \int_{-\infty}^{t} f(t)dt$$

Substituting f(t)dt = dF(t) into the second convolution integral -- a step valid and useful only if f(t) is finite (F(t) continuous) -- we obtain th far more common -- but less useful -- form:

$$\boxed{Y(t) = \int_{-\infty}^{t} X(t-\tau)\, f(\tau)\, dt}$$

These convolution integrals are all originally credited to DUHAM Except in electrical engineering they were little used until the advent of Norbert WIENER. Now due largely to the central role they play in his writi they have come into increasing use throughout engineering and physical scie

Sinusoidal Response and Frequency Domain Behavior

Let us next consider that we disturb a system with an arbitrary input which we characterize in the previous generalized Fourier form, namely:

$$X(t) = \sum_k b_k \cos(\omega_k t + \phi_k)$$

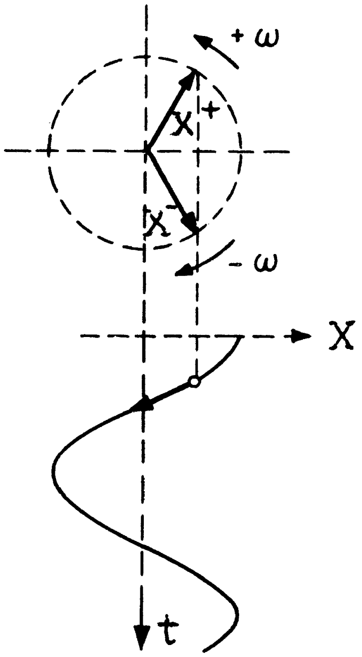
Then if the system is stationary and linear, the response, $Y(t)$, must be that lue to the superposition of responses, Y_k, due to each X_k acting alone. But these individual responses may be derived by examining the behavior of a linear operator excited by a pure sinusoid of frequency ω_k.

Representation of Sinusoids as Phasors

A unit amplitude sinusoidal function of time may always be represented as the instantaneous average of two-counter-rotating vectors (or sinors or phasors) in the form:

$$X(t) = \frac{1}{2}\ [e^{+j\omega t} + e^{-j\omega t}] = \frac{X^+ + X^-}{2}$$

$$\equiv \cos \omega t$$



This principle is actually used s a common type of sinusoidal driver or vibrator.

In this way, we may avoid the artificial use of real and imaginary parts and ay introduce the symmetrical occurrence of positive and negative frequencies.

Frequency Response

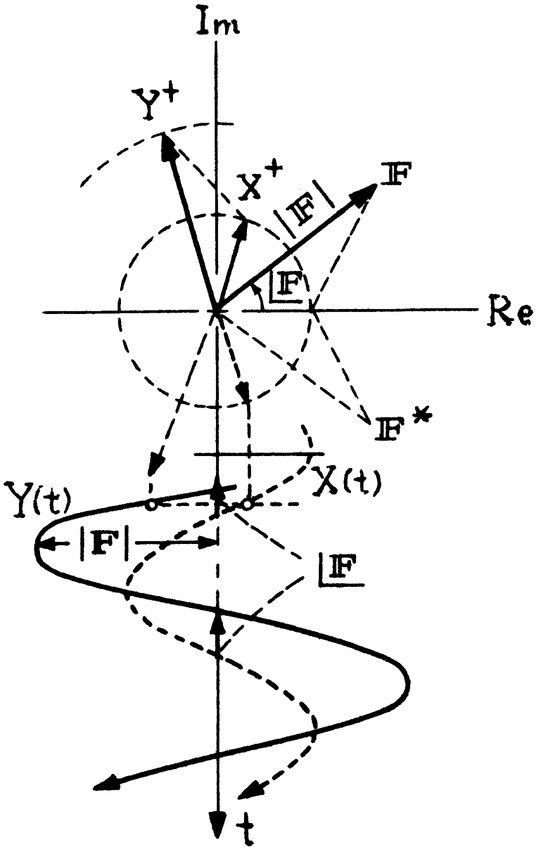
It is then simple to demonstrate that the n-th time derivative of X(t) becomes:

$$D^n X(t) \equiv \frac{1}{2} \left[(+j\omega)^n e^{+j\omega t} + (-j\omega)^n e^{-j\omega t} \right]$$

and therefore that an arbitrary linear operator, $\mathbb{F}(D)$, acting upon X give the result:

$$Y(t) = \mathbb{F}(D)\, X(t) = \frac{1}{2} \left[\mathbb{F}(+j\omega) e^{+j\omega t} + \mathbb{F}(-j\omega) e^{-j\omega t} \right]$$

This may be visualized in terms of phasor diagrams.



It is evident that $\mathbb{F}(-j\omega) = \mathbb{F}^*$ is the complex conjugate of $\mathbb{F}$ just as X^- and Y^- are the instantaneous conjugates of X^+ and Y^+, respectively.

Observe also that the magnitude $|\mathbb{F}|$ and the phase $\underline{|\mathbb{F}}$ have direct interpretation both on the polar and the temporal plots.

Clearly, then, the response of any linear system $\mathbb{F}$ to steady nusoidal excitation is uniquely characterized by the behavior of $\mathbb{F}(j\omega)$, d particularly, $|\mathbb{F}(j\omega)|$ and $\underline{|\mathbb{F}(j\omega)}$. This frequency response may indicated either by the polar locus or Nyquist Plot (in honor of Harry NYQUIST), the magnitude vs phase locus or Nichols Plot (following Nathaniel B. NICHOLS) by the pair of gain (≡ log magnitude) vs frequency and phase vs frequency rves or Bode Plots (after Henryk W. BODE). Certain of these we discuss fur- er below.

plex Frequency Response

Historically, the principal reason for the interest in sinusoidal sponse lay in the characteristic variance of waveform under linear trans- rmations. However, this property is not restricted to sine waves alone; in ct, it also holds for a very natural generalization in the exponentially mped or attenuated sinusoid of the form:

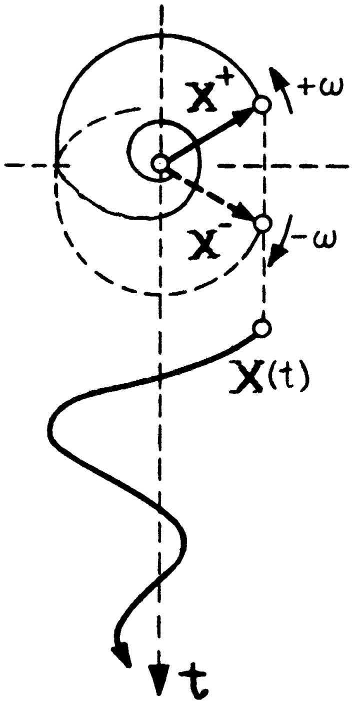
$$X(t) = e^{\sigma t} \cos \omega t$$

t we may now readily generalize as well the previous concept of rotating asors to include this complex form in the fashion:

$$X(t) = \frac{1}{2} [e^{st} + e^{s^*t}]$$

where $s = \sigma + j\omega$

$s^* = \sigma - j\omega$



Such conjugate phasors not only nter-rotate but spirally swell ($\sigma > 0$) decay ($\sigma < 0$) in magnitude as well. us, the complex frequency, s , repre- nts a domain which is simultaneously tating and changing scale, but one in ich the relative phases and magnitudes of phasors remain invariant!

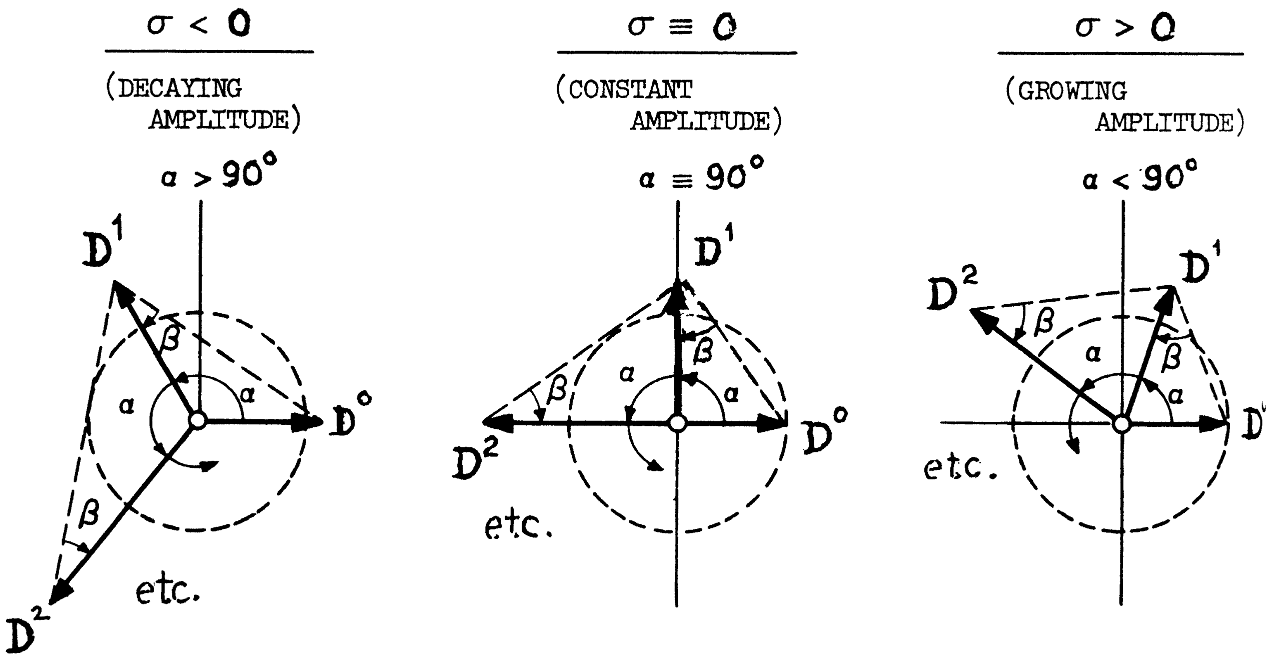
As before we may examine the behavior of $D^n X(t)$. The result is readily obtained as

$$D^n X(t) = \frac{1}{2} [(s)^n e^{st} + (s^*)^n e^{s^*t}]$$

which is equivalent to: $$|D^n| = (\sigma^2 + \omega^2)^{n/2}$$

and: $$\angle D^n = n \tan^{-1} (\omega / \sigma)$$

It is particularly enlightening to interpret these results in phasor form for the three cases:



This representation leads directly to a particularly elegant conceptual picture of the characteristic roots of linear operators simply as those complex values of s for which $1/\mathbb{F}(s) \equiv 0$ and therefore which give vectorial equilibria of the complex phasor diagrams.

The complex frequency response of any linear operator $\mathbb{F}(D)$ is then:

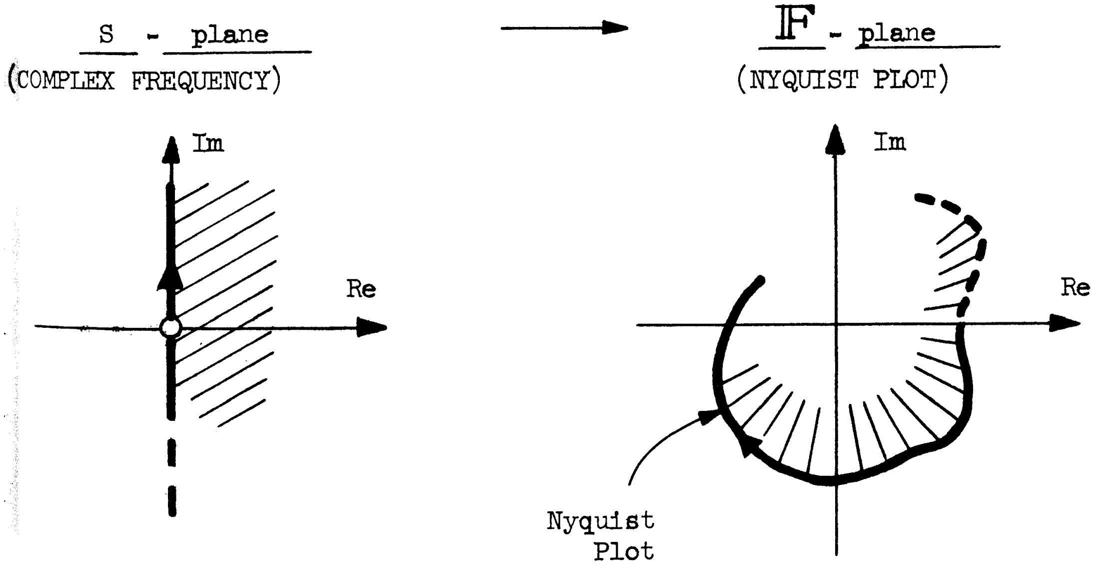
$$Y(t) = \frac{1}{2} [\mathbb{F}(s)e^{st} + \mathbb{F}(s*)e^{s*t}]$$

$$= \frac{1}{2} [\mathbb{F} x^{+} + \mathbb{F}^{*} x^{-}]$$

Again, as with ordinary frequency response, the operator $\mathbb{F}(D)$ is characterized by $\mathbb{F}(s)$.

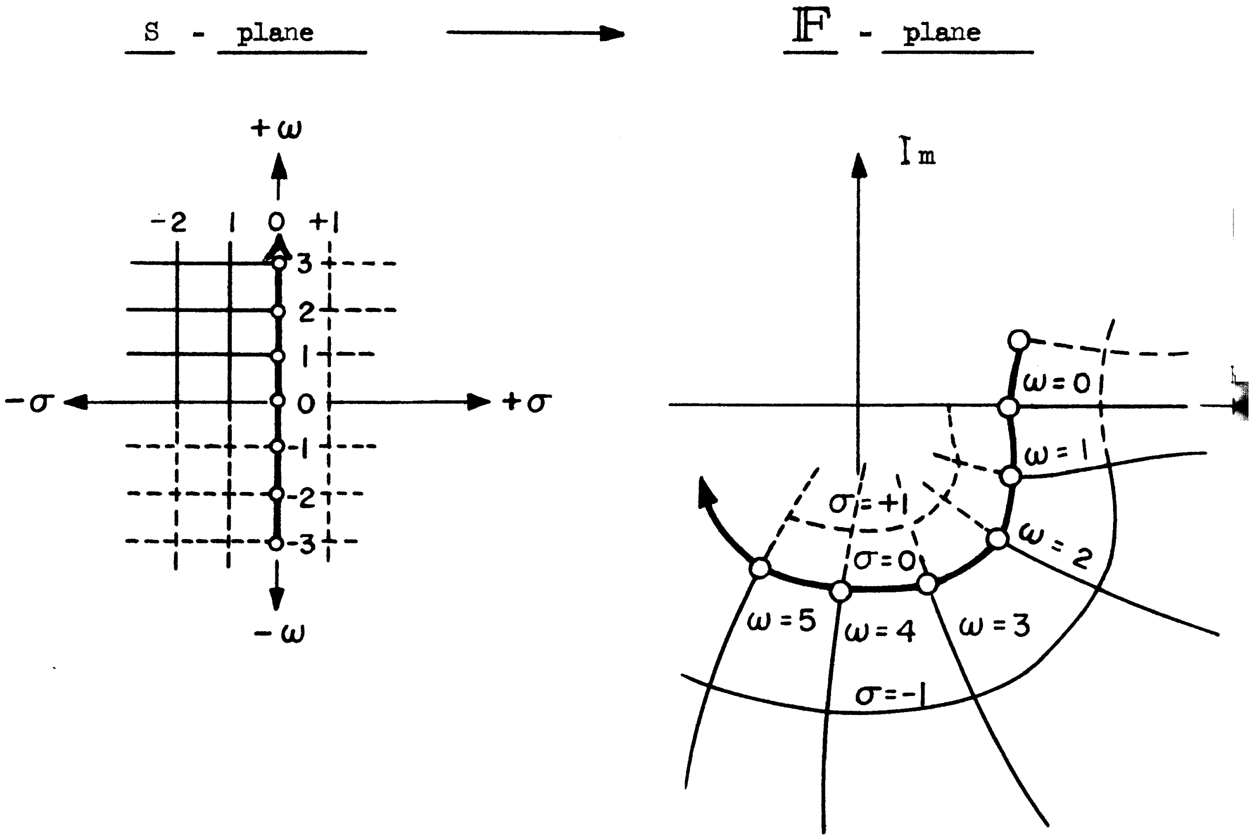
Generalized Frequency Response as a Conformal Mapping Process

The above process can be very effectively interpreted as a conformal mapping process of an s-plane into an $\mathbb{F}$-plane. In particular, the ordinary frequency response becomes a mapping of the imaginary axis $s = j\omega \Big|_{-\infty}^{+\infty}$ into a corresponding curve on the $\mathbb{F}$-plane. This locus is simply the Nyquist Plot in the form:


S - plane (COMPLEX FREQUENCY)

$\mathbb{F}$ - plane (NYQUIST PLOT)

But we now have a much more general context for this mapping, namely the complete complex frequency response in the form:



All the conventional results of frequency response may be obtained from this field map, but useful additional properties may be inferred therefrom. Moreover, this characterization leads directly to the description next following.

H. Linear System Response in terms of Potential Functions

Characterization by Poles and Zeroes

Consider a general system transfer characteristic, $\mathbb{F}(s)$ in terms of the complex root-variable, $s = \sigma + j\omega$. This may generally be expressed as a ratio of polynomials, finite or infinite, in the form:

$$\mathbb{F}(s) = P(s)/Q(s) = \left(\sum_k m_k s^k\right) / \left(\sum_k n_k s^k\right)$$

However, these may in turn be written, at least implicitly, in the factored forms, to give

$$\mathbb{F}(s) = \frac{P(s)}{Q(s)} = \overbrace{A}^{\substack{\text{real}\\ \text{const.}}} \frac{\overset{\text{ZEROES}}{\overset{\downarrow}{(s-p_1)}\,\overset{\downarrow}{(s-p_2)}\cdots}}{\underset{\text{POLES}}{\underset{\uparrow}{(s-q_1)}\,\underset{\uparrow}{(s-q_2)}\cdots}} = A \prod_k \frac{\overset{\downarrow}{(s-p_k)}}{\underset{\uparrow}{(s-q_k)}}$$

The real constant, A, measures the infinite frequency gain of the system, while the roots of the numerator and denominator polynomials give rise to the ZEROES of $\mathbb{F}$ and the POLES of $\mathbb{F}$, respectively. Since $\mathbb{F}$ is completely determined by these terms, one may consider that in this sense, any transfer characteristic may be considered as characterized completely by A and the poles and zeroes.

Potentials of System Functions

If one now divides the transfer characteristic by the constant A and takes the logarithm, there is obtained the normalized transmission characteristic, $\mathbb{G} = \ln(\mathbb{F}/A)$. Thus

$$\mathbb{G}(s) = \ln(\mathbb{F}/A) = \underset{\substack{\uparrow\\ \text{log}\\ \text{magnitude}\\ \text{ratio}}}{\delta(s)} + j\,\underset{\substack{\nwarrow\\ \text{phase}\\ \text{angle}}}{\phi(s)}$$

since $\mathbb{F} = A \cdot \epsilon^{\delta} \cdot \epsilon^{j\phi} = (\text{Mag}) \times \epsilon^{j(\text{phase})}$

By considering the characterization of $\mathbb{F}$ in terms of its poles and zeroes, the transmission characteristic $\mathbb{G}$ can be written:

$$\mathbb{G}(s) = \sum \ln(s - p_k) - \sum \ln(s - q_k)$$

"Sources" ⊕ "Sinks" ⊖

The terms on the right may be considered as the potential function due to a SUM of sources ⊕ and sinks ⊖ located at the zeroes and poles, respectively. This may be visualized and, indeed, realized in an analog consisting of electrical charges, fluid wells, or other situations, thus giving rise to a FIELD in which the equipotentials are contours of constant log magnitude and the fieldlines are contours of constant phase angle.

Log Potential Function for Poles and Zeroes

Consider the function:

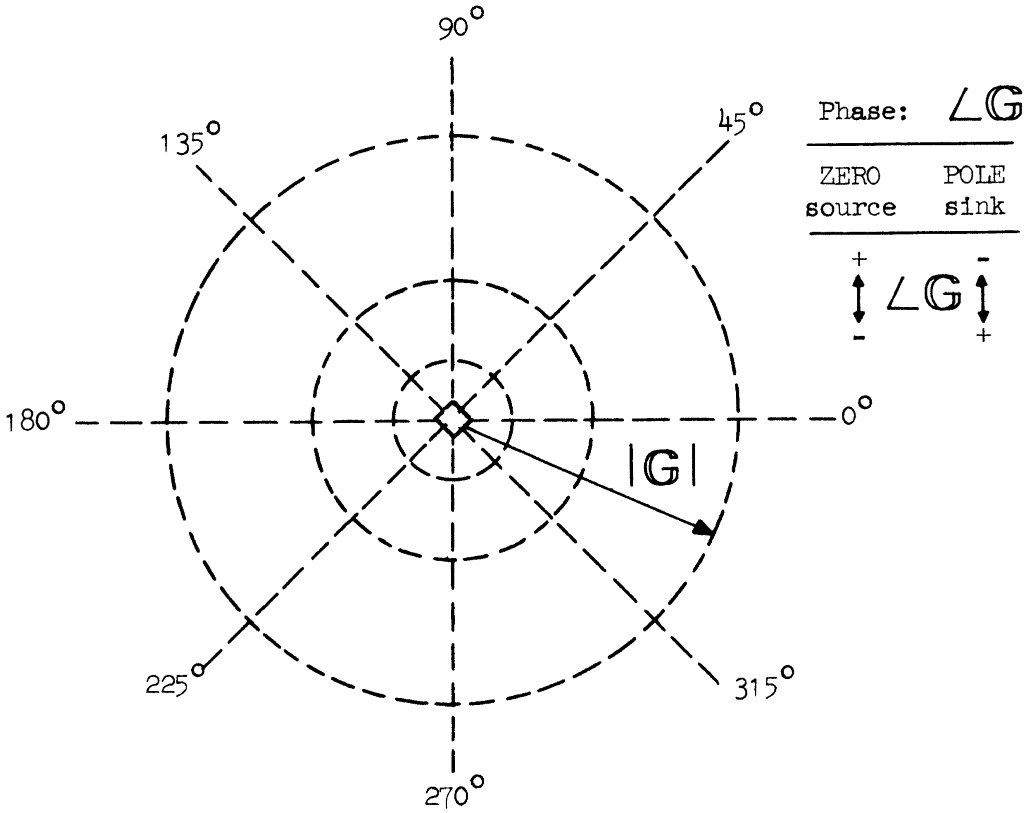
$$\boxed{w(s) = \pm \ln s}$$

+ for ZERO; - for POLE

Now:

RADIAL fieldlines are curves of CONSTANT PHASE

CIRCULAR equipotentials are curves of CONSTANT MAGNITUDE

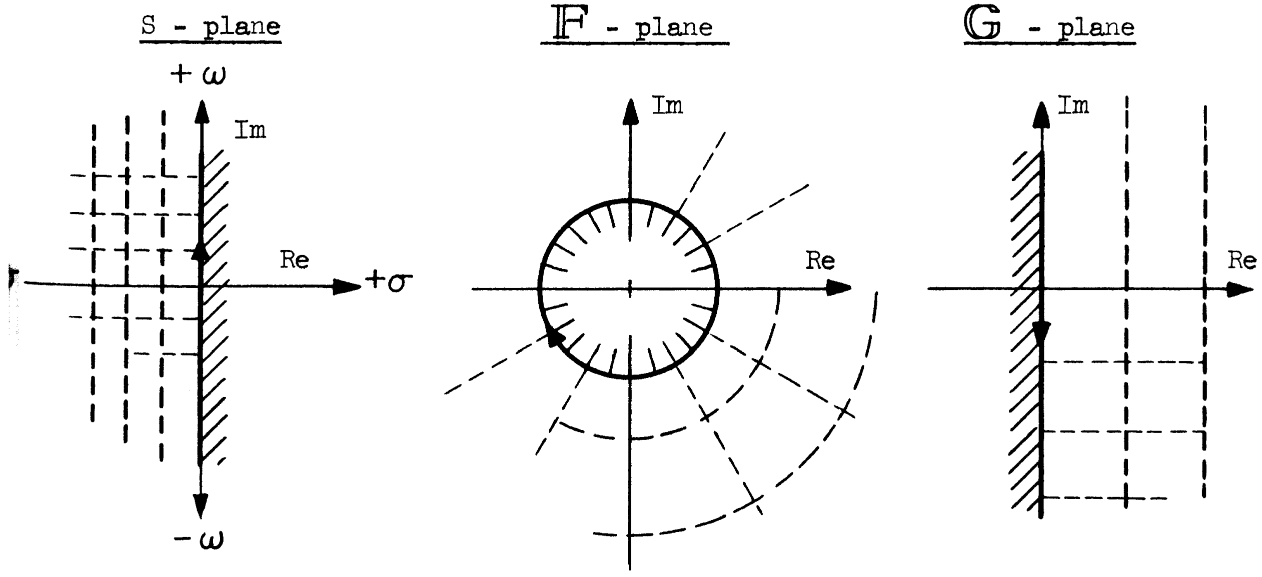


All system functions can be found from this one distribution by locating sources at every pole, sinks at every zero, and then summing up phase angles and log amplitudes for each distribution. This, of course, is identical to vector multiplication of all pole vectors and zero vectors, and is the basis of Walter D. EVANS' "Spirule" and other more complex devices.

Transmission Functions as Potentials

However, in many practical applications of linear system response, it is helpful to return to the basic conception that both $\mathbb{F}(s)$ and $\mathbb{G}(s)$ are analytic functions of the complex frequency $s = \sigma + j\omega$, for which the Cauchy-Riemann equations imply the existence of a set of orthogonal potential functions.

In this light, $\mathbb{G}(s)$ can be viewed as a conformal mapping of $\mathbb{F}(s)$ which is in turn another mapping of the s-plane in the following fashion, for the particular case of a pure delay $\mathbb{F}(s) = e^{-Ts}$:



Clearly in this instance there is neither need for, nor value in, determining the "poles" and "zeroes" of $\mathbb{F}$ before constructing the map of $\mathbb{G}$, since this latter map is even simpler than that for $\mathbb{F}$ itself. Indeed, perhaps the greatest value of the transmission function lies in its generally simple form both for finite reticulations and for continuous systems.

I. One-Port Elements and the Impedance Concept

Generalized One-Port Relations

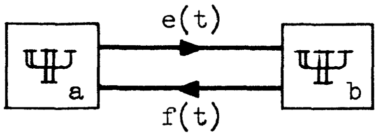
Consider the interaction between two open or closed systems connected by a single energy bond:

$$S_1 \text{———} S_2$$

The causality of this bond could be assigned in only two possible ways:

$$S_1 \vdash\!\!\text{—} S_2 \qquad \text{or} \qquad S_1 \text{—}\!\!\dashv S_2$$

If we now replace each system by a functional operator, Ψ, we find that both cases may be represented by the unique causal scheme:



These two relations, Ψ_a, and Ψ_b, we shall speak of as generalized impedance relations. If both systems are otherwise isolated from the environment, the relations are generally deterministic in nature; if one or both are nonisolated, the operators will necessarily take on stochastic properties. In either event, we recall that the meaning of the functionals Ψ_a and Ψ_b is that the entire history of the input variables (f and e, respectively) is required to establish merely the present value of the output variables (e and f, respectively). Both Ψ functions could represent extremely complex fields, networks, processes, or other systems, but we could still always speak of them as impedance relations, so long as but one port were involved. Some writers, such as KRON, have attempted to generalize further the words impedance and admittance to cover n-port systems where $\mathscr{f}$ and $\mathscr{e}$ vectors were used as inputs, respectively. This usage becomes proportionately clumsier and more specialized as the portality of the elements increases, and/or essential nonlinearities are present.

Therefore we now restrict attention to a deterministic system or subsystem capable of exchanging power only at a single port as indicated above. It is clear that for such elements the over-all behavior is defined by specifying the functional relationship between effort and flow at the single port of entry.

Steady-State Impedance Relations

The static characteristics or steady-state relationships for any one-port element are generally nonlinear static functions of the form:

$$Y = \Phi(X)$$

For the practical (nonideal) case this single curve is usually presented in the form of a graph, defining the range of all possible operating points for the component. This static function may be approximated to an arbitrary degree of precision by a polygonal function (particularly for essentially nonlinear elements), or by an algebraic function (particularly for curvilinear elements).

Dynamic Impedances

From a causal standpoint, since we have seen that the power transfer must depend upon the product of one input variable, $X(t)$, and one output variable $Y(t)$, two general possibilities exist for the nonequilibrium or transient case, namely:

IMPEDANCE FUNCTIONALS:
$X(t) = f(t)$; $Y(t) = e(t)$

$$e(t) = \Psi_{ef} * f(t)$$

ADMITTANCE FUNCTIONALS:
$X(t) = e(t)$; $Y(t) = f(t)$

$$f(t) = \Psi_{fe} * e(t)$$

It is necessary to make a distinction between these two forms, since for the general nonlinear case, a well-defined converse of a given functional relationship may not exist. The choice of the terms, of course, has an historical background.

State-Determined Impedances

The classical state-determined elements [E•, F•, R•, C•, I•] may now all be interpreted as special instances of the generalized impedance functional since

$$
\begin{aligned}
e &= E & & & &\Psi_E \, f \\
e &= \mathbb{R} f &=& & &\Psi_R \, f \\
e &= \mathbb{S} q &=& \ \mathbb{S} \cdot \frac{1}{D} \cdot f &=& \Psi_C \, f
\end{aligned}
$$

$$
\begin{aligned}
f &= F &=& & &\Psi_F \, e \\
f &= \mathbb{G} e &=& & &\Psi_G \, e \\
f &= \Gamma p &=& \ \Gamma \cdot \frac{1}{D} \cdot e &=& \Psi_I \, e
\end{aligned}
$$

History of the Impedance Concept

In electricity the impedance concept grew out of the desire to generalize Ohm's Law and the notion of resistance to make certain elementary direct or constant current concepts applicable to problems involving periodically varying current. Historically, this need arose in the last few decades of the nineteenth century, under pressure of the growing electric power and telephone industries. In dynamic analysis, even before 1900, German and British physical scientists, notable among them HELMHOLTZ, KELVIN, MAXWELL, and HEAVISIDE, saw the analogous structure of electrodynamics and classical dynamics. For these early writers the natural analog of the electrical impedance, relating voltage to current, was the relation of force to velocity. However, largely due to the historical precedence of static elastic analysis in mechanical problems, the principal variables in mechanics were taken to be force and displacement. In this sense, mechanical impedance as a force to displacement ratio may be considered as the attempt to generalize Hooke's Law and the notion of a spring constant to problems in dynamics. In short, the principal justification for differing definitions of impedance between the electrical and mechanical fields is the practical requirement that the steady-state conditions should reduce, at least for small variations, to linear algebraic relations between the principal variables in the respective domains.

In our treatment below we shall loosely use the term impedance to describe all general effort-flow dynamic functional relationships. This renders unnecessary any such distinction between varying definitions.

Besides its roots in classical dynamics, the impedance concept is also firmly founded in the field substructure of material systems, so that the general properties of all energetic one-port elements are derivable from the nature of their underlying fields as outlined elsewhere in these notes. An excellent treatment of the dynamical aspects of the field basis for impedances is given in classical papers by CARSON and by SCHELKUNOFF. The term impedance is credited to BEDELL and CREHORE and its use became widespread through the prolific writings of Charles Proteus STEINMETZ.

Linear One-Port Impedances

If the general functional operators, Ψ_{fe} and Ψ_{ef}, are assumed to be linear operators in the form:

$$\left.\begin{aligned} \mathbb{Z} &\equiv \Psi_{ef} \\ \mathbb{Y} &\equiv \Psi_{fe} \end{aligned}\right\} \text{ LINEAR FORM}$$

then we have the more customary definitions of impedance based upon linear system behavior.

Conventionally, these linear operators, themselves, have been associated with the concept of a linear impedance, namely $\mathbb{Z}$, and its reciprocal, the linear admittance, $\mathbb{Y}$, since now:

$$e = \Psi_{ef} * f = \mathbb{Z} \cdot f$$

$$f = \Psi_{fe} * e = \mathbb{Y} \cdot e$$

$$\text{Thus } \mathbb{Z} \cdot \mathbb{Y} \equiv 1; \quad \text{or } \mathbb{Y} = 1/\mathbb{Z}$$

However, for linear systems with constant (i.e., non-time-varying) parameters, these operators may be advantageously expressed in the form:

$$\mathbb{Z} = Z(D); \qquad \mathbb{Y} = Y(D); \qquad D \equiv d/dt$$

However, an ambiguity in nomenclature would still exist when we attempted to interpret these expressions in causal form. We can preserve our original meaning by interpretating causal impedances and causal admittances in the sense:

$$\text{IMPEDANCE (Relation)} : f \rightarrow e \qquad : \mathbb{Z}$$

$$\text{ADMITTANCE (Relation)} : e \rightarrow f \qquad : \mathbb{Y}$$

as indicated previously.

In any case, all such linear systems become linear one-ports. In the steady-state, a single straight line characteristic will always relate effort to flow.

Background Reading -- Impedance Concept

(1) BEDELL, F. and CREHORE, A.: Derivation and Discussion of the General Solution of the Current Flowing in a Circuit Contianing Resistance, Self-Inductance and Capacity, with Any Impressed Electromotive Force, Journal AIEE, Vol. IX, pp. 303-374 (1892)

(2) CARSON, J. R.: Electromagnetic Theory and the Foundation of Electric Circuit Theory, The Bell System Technical Journal, pp. 1-17. (January, 1927).

(3) SCHELKUNOFF, S. A.: The Impedance Concept and Its Application to Problems of Reflection, Refraction, Shielding and Power Absorption, The Bell System Technical Journal, Vol. 17, pp. 17-48 (January, 1938)

(4) CHENEA, P. F.: On Application of Impedance Method to Continuous Systems, Journal of Applied Mechanics, pp. 571-574 (December, 1953)

F. The Flow of Power and Energy in Systems

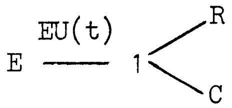
We are now in possession of all the basic tools needed both to answer the questions raised in Part IX concerning the steady-state of energetic systems and also to consider the <u>transient</u> flow of energy over the extent of the system.

We may view this process either in the time domain or in the frequency domain. In the first instance, we merely consider $\mathbb{P}_k(t)$ on each bond k of the system; in the second case $\mathbb{P}_k$ is further spectrally decomposed into $\mathbb{P}_k(\omega_i)$ where the ω_i are all the frequency components present in the power state. Then by a simple generalization of the conventional description employed for a-c power systems, we consider the <u>real</u> power flow, $P_k(\omega_i)$, and <u>reactive</u> power flow, $Q_k(\omega_i)$, along each bond k at frequency i. It may sometimes prove convenient to reticulate further each bond into its spectral components, (P_{ki}, Q_{ki}).

Transient Power and Energy in Reticular Systems

At the beginning of this section we discussed the application of energy principles to state-determined systems, ending with a triangle diagram depicting the history of <u>total</u> system energy for a closed system. Now we may ask how this power and energy is distributed over the extent of the system, recognizing that the microstructure is either continuous or quantized depending upon the ultimate level of energetic reticulation.

It is easiest to demonstrate this state of affairs for a reticular system. Let us take for example the simple R-C system driven by a unit step in effort:



Clearly, we may sensibly inquire as to the values of $[\mathbb{P}_E(t), \mathbb{P}_R(t), \mathbb{P}_C(t)]$ for each of the three bonds, and therefore determine also the corresponding energies $[\mathbb{E}_E(t), \mathbb{E}_R(t), \mathbb{E}_C(t)]$.

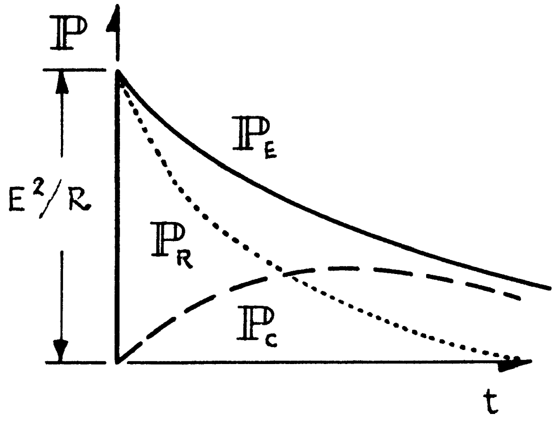
If the system is linear then the following results are readily obtain in terms of $T \equiv RC$:

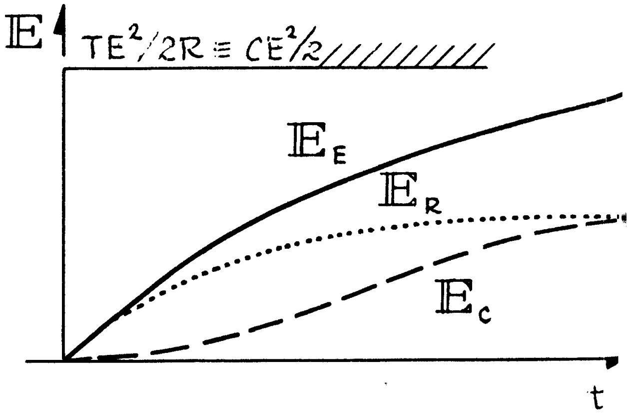
$$\mathbb{P}_E(t) = (E^2/R)\ e^{-t/T} \quad ; \mathbb{E}_E(t) = (TE^2/R)\ [1 - e^{-t/T}\ ]$$

$$\mathbb{P}_R(t) = (E^2/R)\ e^{-2t/T} \quad ; \mathbb{E}_R(t) = (TE^2/2R)\ [1 - e^{-2t/T}\ ]$$

$$\mathbb{P}_C(t) = (E^2/R)\ [e^{-t/T} - e^{-2t/T}\ ] \quad ; \mathbb{E}_C(t) = (TE^2/2R)\ [1 - 2e^{-t/T} + e^{-2t/T}$$

$$\mathbb{P}_E = \mathbb{P}_R + \mathbb{P}_C \quad ; \mathbb{E}_E = \mathbb{E}_R + \mathbb{E}_C$$





Similar results will obtain if $\mathbb{R}$ and $\mathbb{C}$ are nonlinear elements. Moreover, diagrams such as these can be drawn for all bonds of any general reticular system in contact with a particular environment.

If the system is not so reticulated we must return to the field description in terms of a local instantaneous Poynting vector, $\vec{\mathbb{P}}$ (t), and energ density $\mathcal{E}$ (t). However, solutions for such cases are impossible to obtain or yet to conceive except in the simplest classical linear cases.

The Spectral Decomposition of Power into In-Phase and Quadrative Terms

Based upon our remarks above and assuming a generalized Fourier representation for an arbitrary but stationary power state [e(t), f(t)] any single frequency component will have the form:

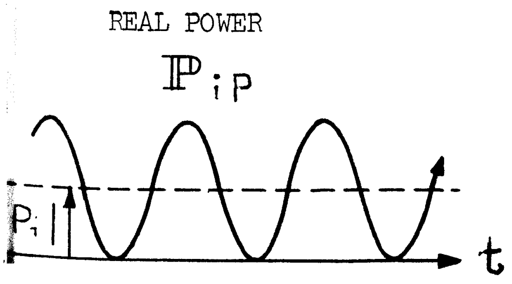
$$e_i = \sqrt{2}\, E_i \cos(\omega_i t)$$

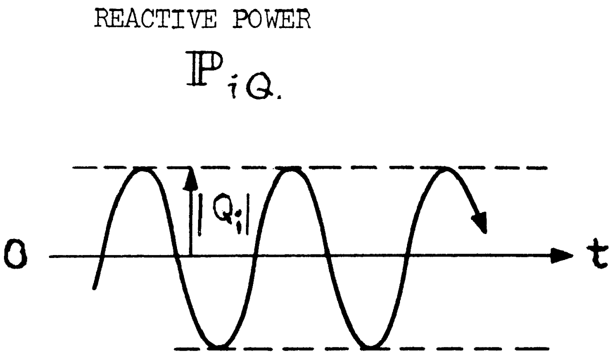
$$f_i = \sqrt{2}\, F_{ip} \cos(\omega_i t) + \sqrt{2}\, F_{iQ} \sin(\omega_i t)$$

Thus the corresponding power component $\mathbb{P}_i \equiv e_i(t) \cdot f_i(t)$ becomes:

$$\begin{aligned}
\mathbb{P}_i &= 2E_iF_{iP} \cos^2(\omega_i t) &+& \quad 2E_iF_{iQ} \cos(\omega_i t)\sin(\omega_i t) \\
&= E_iF_{iP} \,[1 + \cos(2\omega_i t)\,] &+& \quad E_iF_{iQ} \,[\sin(2\omega_i t)\,] \\
&= P_i \,[1 + \cos(2\omega_i t)\,] &+& \quad Q_i \,[\sin(2\omega_i t)\,] \\
&= \mathbb{P}_{iP} &+& \quad \mathbb{P}_{iQ}
\end{aligned}$$

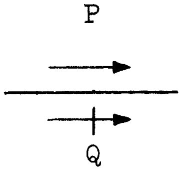
In electrical engineering, the first term has for many years been called the in-phase or real power component and the second term the quadrature or reactive power component. These components have the temporal form:



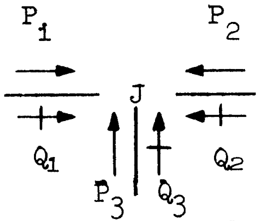


If sgn P_i = sgn Q_i then the reactive power is said to be lagging otherwise it is leading.

In a-c power nets a very useful convention has evolved for indicating the flow of real and reactive power. We may modify this for our purposes as follows for any bond:



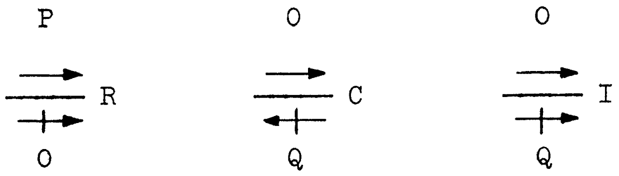
It is then simple to demonstrate that for an ideal energy junction:



$$P_i + P_2 + P_3 \equiv 0$$

$$Q_i + Q_2 + Q_3 \equiv 0$$

Similarly, we may demonstrate for linear elements the following results:



Thus we may apply all the tools of a-c network analysis to the determination of $[P_{ki}, Q_{ki}]$ for each power bond, k , and spectral line, i

In nonlinear devices, such as rectifiers and parametric amplifiers, a number of significant relations exist amongst these doubly reticulated spectral power bonds. Some of these constraints have been studied recently by Paul PENFIELD as generalizations of the earlier MANLEY-ROWE formulas.

Since power may always flow out of a multiport element in a different frequency band than it enters, the macroscopic irreversibility of microscopically reversible systems is no paradox. The generalized forms of the second law of thermodynamics represent attempts to express this intrinsic "band-and-bond" scattering property of all physical systems.

Two-Port Elements and Energy Transport Processes

A. Generalized Two-Port Elements

We have earlier suggested that the behavior of many standard engineering components may be studied by considering them as two-port elements. These two energy ports we shall arbitrarily designate generally as the "upstream" and "downstream" ports, 1 and 2, respectively and are the only ones for which the device is to be represented and investigated. Thus we are led to the problem of specifying the necessary characteristic quantities and relationships to define adequately system behavior, even when the exact internal contents and construction of the two-port system may be unknown or, indeed, "unknowable".

$$\text{PORT 1:} \quad \frac{e_1}{f_1} \quad \text{TWO-PORT} \quad \frac{e_2}{f_2} \quad : \text{PORT 2}$$

In all cases, the behavior will be characterized in terms of the two power flows:

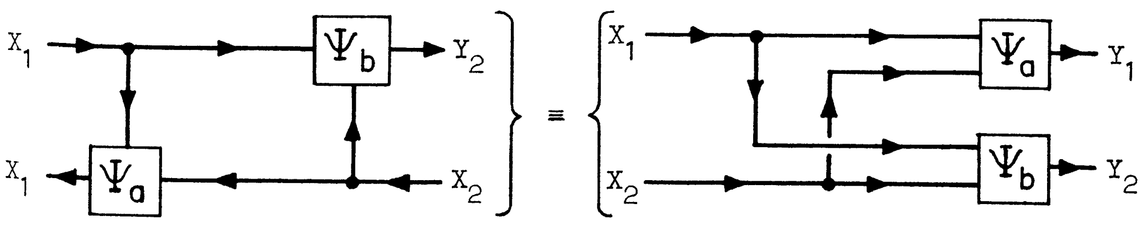
$$\mathbb{P}_1 (t) = e_1 (t) \, . \, f_1 (t); \qquad \mathbb{P}_2 (t) = e_2 (t) \, . \, f_2 (t)$$

and therefore in terms of four variables:

$$e_1 \, , \, f_1 \, , \, e_2 \, , \, f_2$$

Since at each port only one of these variables may be taken as an input, we will generally have the situation depicted below, which indicates that the output variable at each port is functionally determined by the input variables at both ports, that is:

$$Y_1 (t) = \Psi_a [X_1 (t), X_2 (t)]$$
$$Y_2 (t) = \Psi_b [X_1 (t), X_2 (t)]$$



Causally Reticulated Two-Ports

Moreover, only four possible particularizations of this causal sequence exist, namely:

I: IMPEDANCE: ├── 2-Port ──┤

III: ADPEDANCE: ──┤ 2-Port ──┤

II: IMMITTANCE: ├── 2-Port ├──

IV: ADMITTANCE: ──┤ 2-Port ├──

The names employed for the four configurations follow from the partitioning of the words

im - pedance
ad - mittance

and associating the prefixes with the upstream port and the suffixes, with the downstream port of the two-port element. The four possibilities then follow simply from the mnemonic scheme:

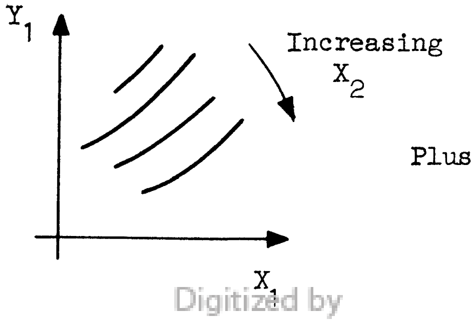
$$\left.\begin{matrix}\text{im}\\ \text{ad}\end{matrix}\right\}\left\{\begin{matrix}\text{pedance}\\ \text{mittance}\end{matrix}\right.$$

Thus we say that a given two-port is in the form of an impedance, we merely imply that the causality of each port, taken alone, is in the form of an impedance functional; by contrast, an adpedance two-port implies that the upstream port has an admittance causality and the downstream port, an impedance causality.

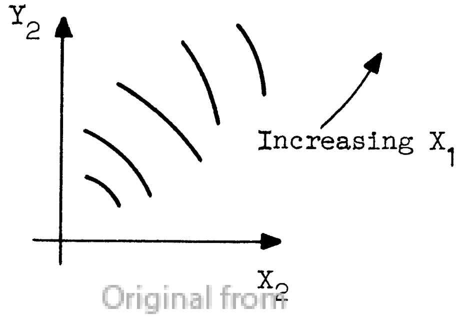
The static characteristics or steady-state relationships for any two-port element are generally nonlinear static functions of the form:

$$Y_1 = \Phi_a(X_1, X_2)$$
$$Y_2 = \Phi_b(X_1, X_2)$$

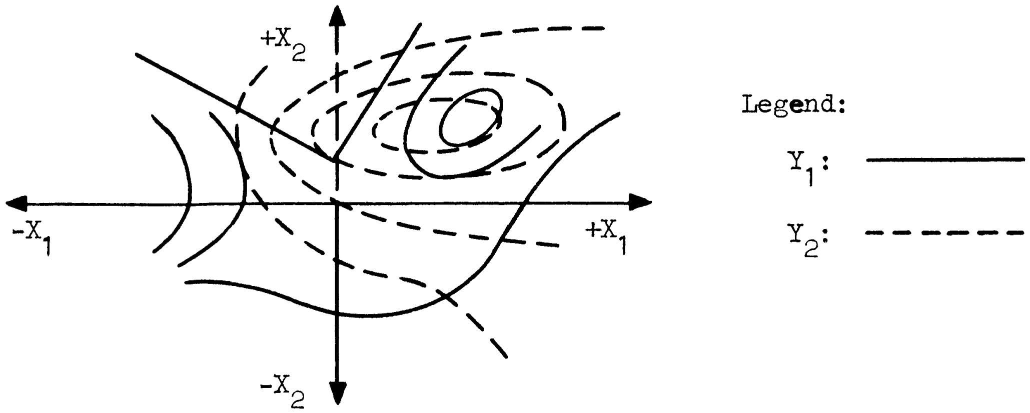
For the general case, these curves are usually depicted in the form of two graphs, each of which relates two of the quantities with the third as a parameter to give a family of curves, as follows:



Plus



However, particularly for operation over both signs of X_1, X_2, the use of "contour" or "hill" characteristics is common; these have the form:



Such static characteristics have the important consequence that they provide a complete specification of the range of "operating points" at which the device or component may be maintained under steady operation, and over which, it may course during transient operation.

B. Primitive Energy Transport Processes

If we consider the typical vehicle propulsion system as indicated below, we can distinctly recognize a small number of basic processes involving the transport of energy from an upstream port to one downstream. The elements used for these purposes may frequently be classified into three primitive types, namely:

a) ENERGY TRANSFORMATION ELEMENTS: [— Transformers —] ≡ [—TF—]
Generalizations of the lever, gear, hydraulic jack, and electrical transformer.

b) ENERGY TRANSDUCTION ELEMENTS: [— Transducers —] ≡ [—TD—]
Generalizations of the motor - and - generator, pump - and - turbine, magnetohydrodynamic devices, heat pump, etc.

c) ENERGY TRANSMISSION ELEMENTS: [— Transmitters —] ≡ [—TM—]
Generalizations of the rod, shaft, pipe, wire, conduit, etc.

Thus the physical scheme:

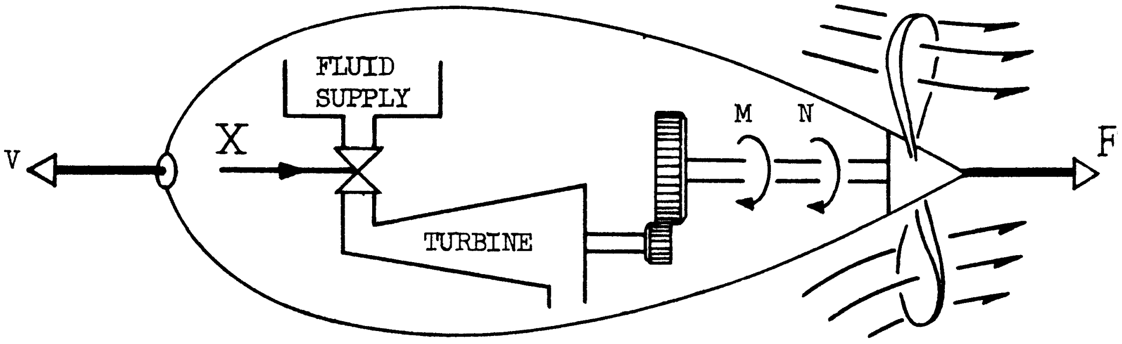
- TURBINE - SHAFT - GEAR - SHAFT - PROPELLER -

Can be represented in general terms by the elements:

-TRANSDUCER-TRANSMITTER-TRANSFORMER-TRANSMITTER-TRANSDUCER-

or ———TD———TM———TF———TM———TD———

The principal benefit of such a generalization is that it readily permits the cataloguing of linear and nonlinear two-port relationships for such components and devices <u>once and for all</u>, quite independently of the media in which the devices operate.



VEHICLE PROPULSION SYSTEM

C. Linear Two-Port Elements

If the general functional operators for the two-port, Ψ_a, and Ψ_b, can be presumed linear over the practical range of operation of a particular component, then a most significant and powerful simplification subsists. This is manifested in the reductions:

$$Y_1 = \Psi_a * [X_1, X_2] = \mathbb{F}_{11} \cdot X_1 + \mathbb{F}_{12} \cdot X_2$$

$$Y_2 = \Psi_b * [X_1, X_2] = \mathbb{F}_{21} \cdot X_1 + \mathbb{F}_{22} \cdot X_2$$

which may then be summarized in the single causal statement, in matrix form:

$$\left[\begin{array}{c} Y_1 \\ \hdashline Y_2 \end{array}\right] \equiv \mathbb{Y} = \Lambda \cdot \mathbb{X} = \left[\begin{array}{c:c} \mathbb{F}_{11} & \mathbb{F}_{12} \\ \hdashline \mathbb{F}_{21} & \mathbb{F}_{22} \end{array}\right] \cdot \left[\begin{array}{c} X_1 \\ \hdashline X_2 \end{array}\right]$$

Causal Matrices

The matrix Λ then has the four physically realizable particularizations; these have evolved into a fairly standard symbolism over the last several decades, at least as regards the electrical field. The four causal matrices, in standard forms, are as follows:

Configuration I ....Impedance Matrix: $\mathbb{Z} \equiv \left[\begin{array}{c:c} z_{11} & z_{12} \\ \hdashline z_{21} & z_{22} \end{array}\right]$

Configuration II ....Immittance Matrix: $\mathbb{H} \equiv \left[\begin{array}{c:c} h_{11} & h_{12} \\ \hdashline h_{21} & h_{22} \end{array}\right]$

Configuration III ....Adpedance Matrix : $\mathbb{G} \equiv \left[\begin{array}{c:c} g_{11} & g_{12} \\ \hdashline g_{21} & g_{22} \end{array}\right]$

Configuration IV ....Admittance Matrix: $\mathbb{Y} \equiv \left[\begin{array}{c:c} y_{11} & y_{12} \\ \hdashline y_{21} & y_{22} \end{array}\right]$

The relation of these causal matrices to the possible combinations of 2-ports will be indicated subsequently.

Transmission Matrices:

In addition to these four causal matrices a most significant standard form was long ago developed, which established a direct spatial correspondence to the ports themselves. This relates the power states, $\mathbb{S}_1$ and $\mathbb{S}_2$, at each end of the linear two-port through a transmission matrix in the form:

$$\mathbb{S}_1 = \mathbb{M} \cdot \mathbb{S}_2$$

$$\left[\begin{array}{c} e_1 \\ \hline f_1 \end{array}\right] = \left[\begin{array}{c|c} \mathbb{A} & \mathbb{B} \\ \hline \mathbb{C} & \mathbb{D} \end{array}\right] \cdot \left[\begin{array}{c} e_2 \\ \hline f_2 \end{array}\right]$$

Of course, the inverse transmission matrix, $\mathbb{M}^{-1}$, relates $\mathbb{S}_2$ to $\mathbb{S}_1$ in the form:

$$\mathbb{S}_2 = \mathbb{M}^{-1} \mathbb{S}_1$$

While the $\mathbb{M}$ and $\mathbb{M}^{-1}$ matrices are clearly noncausal, they have the peculiar advantage that the overall coupled transmission matrix for two 2-ports in tandem or cascade may be obtained by direct matrix multiplication in the fashion:

$$\left[\begin{array}{c} e_1 \\ \hline f_1 \end{array}\right] = \left[\begin{array}{c|c} \mathbb{A}_a & \mathbb{B}_a \\ \hline \mathbb{C}_a & \mathbb{D}_a \end{array}\right] \cdot \left[\begin{array}{c|c} \mathbb{A}_b & \mathbb{B}_b \\ \hline \mathbb{C}_b & \mathbb{D}_b \end{array}\right] \cdot \left[\begin{array}{c} e_2 \\ \hline f_2 \end{array}\right]$$

$$\left[\begin{array}{c} e_1 \\ \hline f_1 \end{array}\right] = \left[\begin{array}{c|c} \mathbb{A}_a\mathbb{A}_b + \mathbb{B}_a\mathbb{C}_b & \mathbb{A}_a\mathbb{B}_b + \mathbb{B}_a\mathbb{D}_b \\ \hline \mathbb{C}_a\mathbb{A}_b + \mathbb{D}_a\mathbb{C}_b & \mathbb{C}_a\mathbb{B}_b + \mathbb{D}_a\mathbb{D}_b \end{array}\right] \cdot \left[\begin{array}{c} e_2 \\ \hline f_2 \end{array}\right]$$

Historical Notes

A long history is associated with the development of linear two-port concepts. In electrical engineering these systems have been known by the various alternative names:

FOUR TERMINAL NETWORK
TWO-TERMINAL-PAIR NETWORK
QUADRUPOLE (or QUADRIPOLE)
FOURPOLE (or "VIERPOL" in German)

as well as a number of others. The designation fourpole was first used by Breisig in 1921. While the first practical use of such concepts was in the theory of long power transmission lines and associated apparatus including transformers, Breisig apparently was the first to have used the $\mathbb{A}$, $\mathbb{B}$, $\mathbb{C}$, $\mathbb{D}$, operators for communication lines and networks. The application of the matrix notation is credited to Strecker and Feldtkeller. Much additional valuable material on general linear two-ports can be found in works by Baerwald, Guillemin, Pipes, Le Corbeiller, and many others.

Origin of the Term Port

The word "port" used in this connection apparently originated with Harold A. WHEELER to describe the coupling holes in waveguides. This usage was presented by Wheeler to the IRE in the following words:

> It has been customary to designate each entrance or exit of a network as a pair of terminals, based upon the circuit concept of wires and conduction. The result was cumbersome terms such as "four-terminal pair" with the unobvious meaning of a network with two pairs of terminals. Furthermore, the terminal-pair concept becomes artificial in the case of electromagnetic fields transmitting power within boundaries, through holes, and from one region to another in space.
>
> After considering many alternatives the writer has adopted the term ... "port" as the general designation of an entrance or exit of a network. A self-impedance becomes a "one-port". The usual transducer becomes a "two-port" ... The general network ... a "multi-port". This plan ... is first put to use in this monograph.

Reciprocity and Symmetry

Certain special relationships frequently exist for linear two-ports which may be directly expressed in terms of the matrix elements, namely:

$$\text{PASSIVE RECIPROCITY:} \ldots\ldots \begin{cases} Z_{12} = Z_{21};\ Y_{22} = Y_{21} \\ \det \mathbb{M} \equiv \Delta = \mathbb{A}\,\mathbb{D} - \mathbb{B}\,\mathbb{C} = 1 \end{cases}$$

$$\text{SPACE SYMMETRY:} \ldots\ldots\ldots \quad \mathbb{A} \equiv \mathbb{D}$$

Thus any symmetrical, reciprocal (i.e., PASSIVE) linear two-port can be described completely in terms of only two linear operators, for example, $\mathbb{A}$ and $\mathbb{B}$ since

$$\mathbb{M} = \left[\begin{array}{c:c} \mathbb{A} & \mathbb{B} \\ \hdashline (\mathbb{A}^2-1)/\mathbb{B} & \mathbb{A} \end{array}\right] \quad \text{for} \left\{\begin{array}{c} \text{RECIPROCAL} \\ \text{and} \\ \text{SYMMETRIC} \end{array}\right\} \text{---} \quad \text{LINEAR TWO PORTS}$$

Two-Ports Composed of 1-Port Impedances and 3-Port Junctions

An extensive and useful subclass of two-port elements arises from the interconnection of one-port elements, coupled through energy junctions, in a polymerized chain of the form:

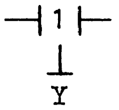
$$[\;—\underset{\substack{\bullet\\ Z}}{J}—\underset{\substack{\bullet\\ Z}}{J}—\cdots—\underset{\substack{\bullet\\ Z}}{J}—\;]$$

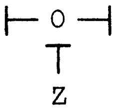
where
J = 0 or 1
Z = 1-Port Impedances

Since the individual one-ports could be either impedance or admittance functionals, both forms of energy junction are involved and two different particularizations exist:

$$\overbrace{—\underset{\substack{\bullet\\ Z}}{1}— \qquad \text{and} \qquad —\underset{\substack{\bullet\\ Y}}{0}—}^{—\underset{\substack{\bullet\\ Z}}{J}—}$$

We could call these structures impedance two-ports and admittance two-ports in strict conformance with our general usage. However, we should note that since the conventional causality for each junction would give the forms:



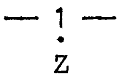


there results a "cross-over" between these causal usages and the conventional noncausal usage immediately following. The best solution seems to be to call the structure with the effort junction a SERIES IMPEDANCE and that with the flow junction, a SHUNT ADMITTANCE, which agrees with standard electrical terminology.

Linear Impedance and Admittance 2-Ports

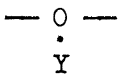
For linear 1-ports the corresponding series impedance and shunt admittance 2-port transmission matrices become:

SERIES IMPEDANCE



$$\left[\begin{array}{c:c} 1 & \mathbb{Z} \\ \hdashline 0 & 1 \end{array}\right]$$

SHUNT ADMITTANCE

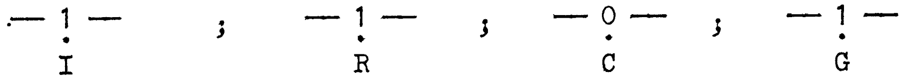


$$\left[\begin{array}{c:c} 1 & 0 \\ \hdashline \mathbb{Y} & 1 \end{array}\right]$$

From our previous analysis it should now be amply clear that both $\mathbb{Z}$ and $\mathbb{Y}$ might consist of any number of energy storage and dissipating elements, so long as only one energy port communicates with the rest of the system. However, for each matrix, two simple cases are particularly noteworthy; namely, where:

$$\mathbb{Z} = \begin{cases} \mathrm{ID} \\ \mathrm{or} \\ \mathrm{R} \end{cases} \quad ; \quad \mathbb{Y} = \begin{cases} \mathrm{CD} \\ \mathrm{or} \\ \mathrm{G} \end{cases}$$

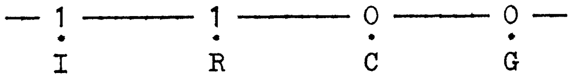
We may indicate these by bond graphs and corresponding matrices:



$$\left[\begin{array}{c|c}1 & ID\\ \hline 0 & 1\end{array}\right] ; \quad \left[\begin{array}{c|c}1 & R\\ \hline 0 & 1\end{array}\right] ; \quad \left[\begin{array}{c|c}1 & 0\\ \hline CD & 1\end{array}\right] ; \quad \left[\begin{array}{c|c}1 & 0\\ \hline G & 1\end{array}\right]$$

Examples of Polymerized Chains

Using the above elements, we may now demonstrate their unusual value for rapid determination of 2-port characteristics of typical polymers For example, a commonly encountered form of damped oscillator in any medium has the following properties:

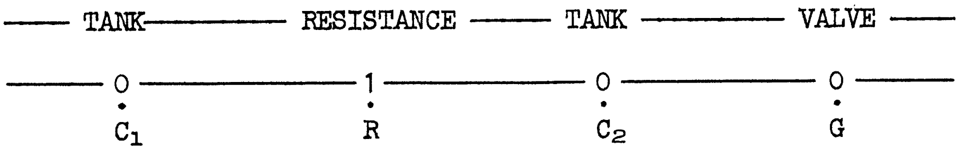


$$\left[\begin{array}{c|c}1 & ID\\ \hline 0 & 1\end{array}\right] \left[\begin{array}{c|c}1 & R\\ \hline 0 & 1\end{array}\right] \left[\begin{array}{c|c}1 & 0\\ \hline CD & 1\end{array}\right] \left[\begin{array}{c|c}1 & 0\\ \hline G & 1\end{array}\right]$$

$$\left[\begin{array}{c|c}1 & ID + R\\ \hline 0 & 1\end{array}\right] \left[\begin{array}{c|c}1 & 0\\ \hline CD + G & 1\end{array}\right]$$

$$\left[\begin{array}{c|c}1+(ID+R)(CD+G) & ID + R\\ \hline CD + G & 1\end{array}\right]$$

In a similar fashion we may depict a certain linearized fluid system in the form:



$$\left[\begin{array}{c|c}1 & 0\\ \hline C_1D & 1\end{array}\right] \left[\begin{array}{c|c}1 & R\\ \hline 0 & 1\end{array}\right] \left[\begin{array}{c|c}1 & 0\\ \hline C_2D & 1\end{array}\right] \left[\begin{array}{c|c}1 & 0\\ \hline G & 1\end{array}\right]$$

$$\left[\begin{array}{c|c}1 + RG + T_{12}D & R\\ \hline G[1+(T_{11}+T_{21}+T_{22})D+T_{11}T_{22}D^2] & 1 + T_{11}D\end{array}\right]$$

where: $T_{11} \equiv RC_1$; $T_{12} = RC_2$; $T_{21} = C_1/G$; $T_{22} = C_2/G$

D. Some Standard Forms of Two-Port Nets:

A number of recurring 2-port structures formed from impedance functionals have been given names in electrical science. As mentioned previously, the topology of such nets can be described simply in terms of the junction structure alone, using the conventions:

$$[\,-\,\underset{\bullet}{1}\,-\,] \;=\; [\,-\,\underset{\underset{Z}{\bullet}}{1}\,-\,] \;\equiv\; [\,-\,ZZ\,-\,]$$

$$[\,-\,\underset{\bullet}{0}\,-\,] \;=\; [\,-\,\underset{\underset{Y}{\bullet}}{0}\,-\,] \;\equiv\; [\,-\,YY\,-\,]$$

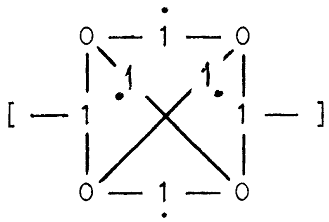
The canonical structures may then be enumerated in a unique simple order as follows:

1) The " L - Net " : $[\,-\,\underset{\bullet}{0}\,-\,\underset{\bullet}{1}\,-\quad\,] \equiv [\,-\,EL\,-\,]$

2) The " PI - Net " : $[\,-\,\underset{\bullet}{0}\,-\,\underset{\bullet}{1}\,-\,\underset{\bullet}{0}\,-\,] \equiv [\,-\,PI\,-\,]$

3) The "Tee - Net" : $[\,-\,\underset{\bullet}{1}\,-\,\underset{\bullet}{0}\,-\,\underset{\bullet}{1}\,-\,] \equiv [\,-\,TEE\,-\,]$

4) The "Lattice - Net" or "Bridge - Net" :



$\equiv [\,-\,LAT\,-\,]$

These structures may in turn be cascaded or polymerized to form the following ladder nets:

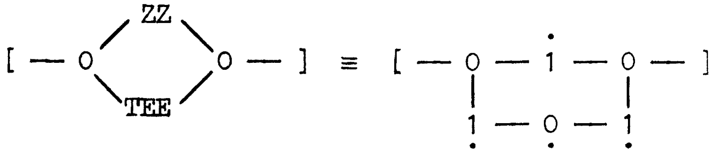
5) The " L - Ladder " $[\,-\,EL\,-\,]^n \;=\; [\,-\underset{\cdot}{0}-\underset{\cdot}{1}-\underset{\cdot}{0}-\underset{\cdot}{1}- \quad -\underset{\cdot}{0}-\underset{\cdot}{1}-\quad]$

6) The " Pi - Ladder " $[\,-\,PI\,-\,]^n \;=\; [\,-\underset{\cdot}{0}-\underset{\cdot}{1}-\underset{\cdot}{0}-\underset{\cdot}{1}- \quad -\underset{\cdot}{0}-\underset{\cdot}{1}-\underset{\cdot}{0}-]$

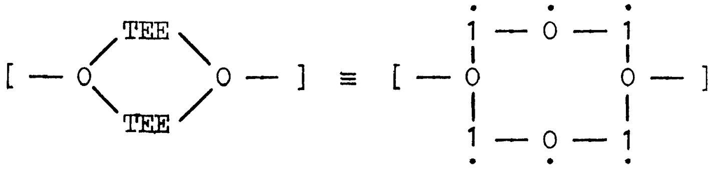
7) The "Tee-Ladder" $[\,-\,TEE\,-]^n \;=\; [\,-\underset{\cdot}{1}-\underset{\cdot}{0}-\underset{\cdot}{1}-\underset{\cdot}{0}- \quad -\underset{\cdot}{1}-\underset{\cdot}{0}-\underset{\cdot}{1}-]$

and may be combined in various other ways to form other named structures such as

8) The "Bridged - T"



9) The "Parallel - T"



The above topologies may all be represented uniquely in coded form as follows:

a) <u>ZZ</u> ab ; 1 ab.

b) <u>YY</u> ab ; 0 ab.

1) <u>EL</u> ab ; YY ac ; ZZ bc.

2) <u>PI</u> ab ; YY ac ; ZZ cd ; YY bd.

3) <u>TEE</u> ab ; ZZ ac ; YY cd ; ZZ bd.

4) <u>LAT</u> ab ; 1 acd ; 1 bef ; 0 cgh ; 0 dij ; 0 ekl ; 0 fm
0 fmn ; ZZ gk ; ZZ jn ; ZZ hm ; ZZ il .

8) <u>Bridged</u> <u>TEE</u> ab; 0 acd ; 0 bef ; ZZ ce ; TEE df.

9) <u>Parallel</u> <u>TEE</u> ab; 0 acd ; 0 bef ; TEE ce ; TEE df.

E. Description of Linear Two-Ports

Each of the six forms of matrix for the linear two-port (i.e. $\mathbb{Z}$, $\mathbb{G}$, $\mathbb{H}$, $\mathbb{Y}$, $\mathbb{M}$, $\mathbb{M}^{-1}$) involves for the general case four independent functional operators, which most simply are the corresponding four matrix elements, themselves. The specification of these four elements completely determines the behavior of a linear two-port; in particular, at any given frequency four suitable measurements will suffice to describe response.

If the network is passive reciprocal, only three of the functional operators are independent and there is thus one constraining relation among the set of four; now, at any single frequency three measurements will suffice to define the system.

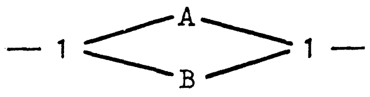
These various methods of specifying system behavior are interrelated so that given one set we may find any of the others. We shall here consider these relations.

The Interconnection of Two-Ports

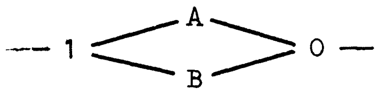
There are five possible ways of interconnecting a pair of two-port elements; these may be described as follows:

1) Cascade:

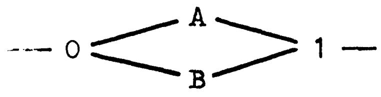
2) Series - Series:



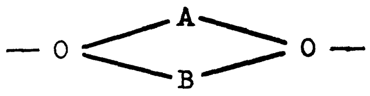
3) Series - Shunt:



4) Shunt - Series:



5) Shunt - Shunt:



In the cascade connection we have already seen that the $\mathbb{M}$ - matrices simply multiply to obtain the interconnected behavior. The remaining connections are governed by the $\mathbb{Z}$, $\mathbb{H}$, $\mathbb{G}$, $\mathbb{Y}$ matrices respectively, as follows:

2) Series - Series: $\mathbb{Z} = \mathbb{Z}_a + \mathbb{Z}_b$
3) Series - Shunt : $\mathbb{H} = \mathbb{H}_a + \mathbb{H}_b$
4) Shunt - Series: $\mathbb{G} = \mathbb{G}_a + \mathbb{G}_b$
5) Shunt - Shunt: $\mathbb{Y} = \mathbb{Y}_a + \mathbb{Y}_b$

The extension of these results to *systems* of two-ports cascaded and coupled through energy junctions is straightforward.

Transfer Characteristics of Two-Port Nets

The questions which arise in connection with systems involving linear two-ports are commonly grouped in the following categories:

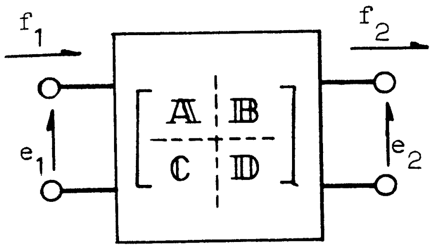
I. The TRANSFER Problem: wherein one seeks the effort or flow at the downstream port in response to effort or flow at the upstream port, with *ideal* terminations generally assumed at the downstream port;

II. The TRANSMISSION Problem: wherein the power state at one port is required in terms of the power state at the second port with:
  a) Unrestricted terminal conditions, or
  b) Terminal impedances specified;

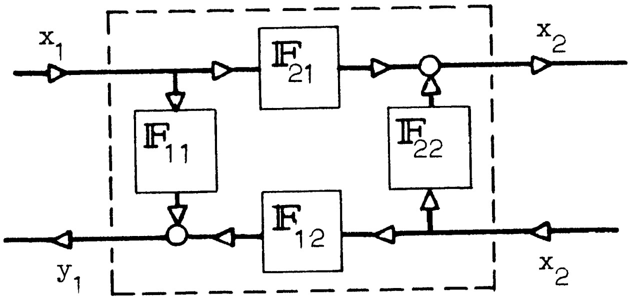
III. The INSERTION Problem: wherein is sought the effect of inserting a two-port into a system in place of a through bond. Typically these problems are "filtering" and "protection" situations, where performance is measured in terms of the *change* in power, effort, or flow after insertion from that occurring before insertion.

All these problems require consideration of the *transfer* characteristics of a given two-port. These may be readily determined in terms of the elements of the $\mathbb{M}$ - matrix as indicated in the attached table.

MATRIX FORM:



CAUSAL FORM:



Determinant $\Delta = \begin{vmatrix} \mathbb{A} & \mathbb{B} \\ \mathbb{C} & \mathbb{D} \end{vmatrix} \equiv \mathbb{A}\mathbb{D} - \mathbb{B}\mathbb{C}$

$\cong$ 1 for PASSIVE (RECIPROCAL) SYSTEMS

CASE	VARIABLES				TRANSFER OPERATORS			
	Inputs		Outputs		$\mathbb{F}_{11}$	$\mathbb{F}_{12}$	$\mathbb{F}_{21}$	$\mathbb{F}_{22}$
	x_1	x_2	y_1	y_2	$\partial y_1/\partial x_1$	$\partial y_1/\partial x_2$	$\partial y_2/\partial x_1$	$\partial y_2/\partial x_2$
$\mathbb{Y}$	e_1	e_2	f_1	f_2	$\mathbb{D}/\mathbb{B}$	$-\Delta/\mathbb{B}$	$1/\mathbb{B}$	$-\mathbb{A}/\mathbb{B}$
$\mathbb{G}$	e_1	f_2	f_1	e_2	$\mathbb{C}/\mathbb{A}$	$+\Delta/\mathbb{A}$	$1/\mathbb{A}$	$-\mathbb{B}/\mathbb{A}$
$\mathbb{H}$	f_1	e_2	e_1	f_2	$\mathbb{B}/\mathbb{D}$	$+\Delta/\mathbb{D}$	$1/\mathbb{D}$	$-\mathbb{C}/\mathbb{D}$
$\mathbb{Z}$	f_1	f_2	e_1	e_2	$\mathbb{A}/\mathbb{C}$	$-\Delta/\mathbb{C}$	$1/\mathbb{C}$	$-\mathbb{D}/\mathbb{C}$

TRANSFER OPERATORS FOR LINEAR TWO-PORTS

Background Reading -- Early 2.Port Literature

(1) EVANS, R. D. and SELS, H. K.: Transmission Line Constants and Resonance, Electrical Journal, Vol. 18, p. 306 (1921)

Introduced ABCD constants to power engineers.

(2) BREISIG, F.: Theoretische Telegraphie, Second Edition (1924)

(3) STRECKER, F. and FELDTKELLER, R.: Grundlagen der Theorie des allgemeinen Vierpols, Elektrische Nachrichten Tecknik, Vol. 6, p. 93 (1929)

(4) BAERWALD, H. G.: Die Eigenschaften Symmetrischer 4n-Pol..., Sitzb. d. Preuss Akad. d. Wiss, Phys-Math. Kl., Vol. 33, p. 784 (19

The above three works carried over the use of linear 2.port concepts into communications circuits.

(5) WAGNER, K. W.: Operatorenrechnung (1940)

Related the 2.port operators to transient response.

Background Reading -- Early Use of 2.port Matrices

(1) GUILLEMIN, E. A.: Communication Networks, Vol. II (1935)

(2) PIPES, L. A.: The Matrix Theory of Four Terminal Networks, Phil. Mag. Vol. 30, p. 370 (1940).

These two authors are principally responsible for the present nearly universal use of matrices for linear system analysis.

Background Reading -- Current Works

(1) Le CORBEILLER, P.: Matrix Analysis of Electrical Networks (1950)

A very readable account of the causal 2.port matrices.

(2) WEBER, Ernst: Linear Transient Analysis, Vol. II (1956)

Nearly the entire volume is devoted to linear passive and active 2.ports largely treated in terms of transmission matrices.

(3) KUPFMULLER, K.: Einfuhrung in die theoretische Electrotechnik, Fifth Edition (1957).

This classic German text has been kept up-to-date and matrix methods are used throughout.

(4) LAURENT, Torbern: Vierpoltheorie und Frequenztransformation (1956)

A book principally concerned with frequency domain behavior for communication systems.

Background Reading -- Gyrators

(1) TELEGEN, B. D. H.: The Gyrator, a New Electric Network Element, Philips Research Reports, p. 81, (April, 1948).

This epic-making paper has stirred up a revolution in microwave techniques.

XV. Transformers and Transducers

A. The Concept of Ideal Two-Port Elements

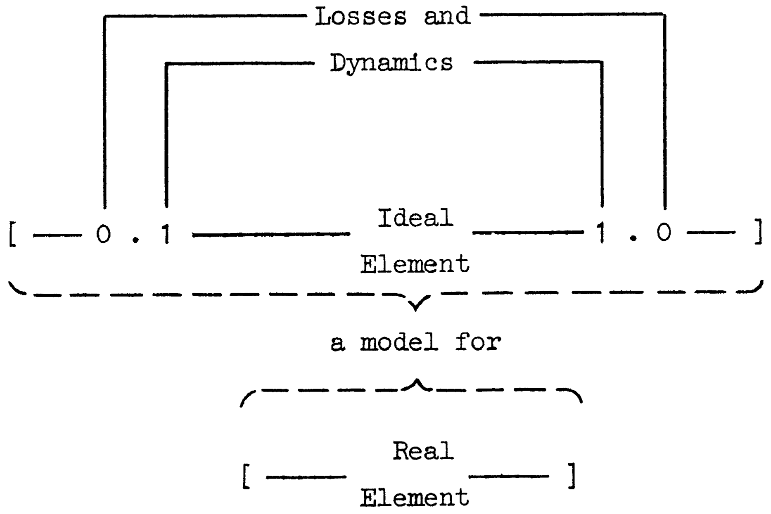
We have earlier introduced the concept of a class of ideal multiport elements for which

$$\sum \mathbb{P} \equiv 0$$

where the sum is carried over all ports. In the particular case of three ports, the ideal effort and flow junctions satisfied this condition. We shall now consider two-port elements of this class indicated as follows:

[——Ideal Element——] or [——IE——]

Losses of effort and flow, as well as dynamic effects, such as inertance and capacitance, can then be appended to these ideal elements to model certain types of real elements, in the fashion:



Particularly for energy transducers this need for "interdiffusion" of media exists; it is not generally possible to represent the real process of energy conversion without energy losses which involve state variables from both media concerned.

However, for many energy transformers, it is often possible to embed the losses and dynamics in <u>transmission elements</u> at the two ports in the fashion:

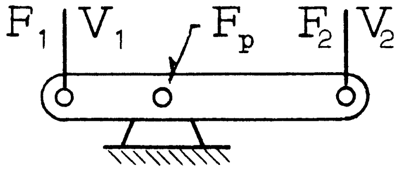
— Transmission Element —— Ideal Element —— Transmission Element

or: — TM —— IE —— TM —

The use of the above reticulations for practical representations involving the transformation and transduction of power we shall now discuss.

B. Energy Transformation Elements [—— TF ——]

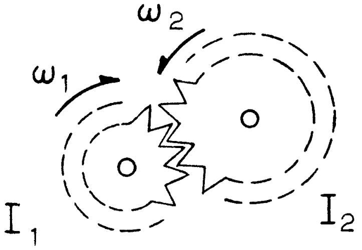
These useful devices may always be considered in a lossless static form, since dynamics and dissipation can be readily included immediately upstream and downstream of the transformer ports. This may be seen readily in terms of particular instances. Consider, for example, the treatment of pivot friction in a lever:



$$\Bigg\} \equiv [\,-\,\underset{\substack{\cdot\\ R}}{1}\,-\,TF\,-\,] \text{ or } [\,-\,TF\,-\,\underset{\substack{\cdot\\ R}}{1}\,-\,]$$

SCHEMATIC MOLECULES

Similarly, the inertia of a real set of spur gears could be appended to the terminals of the primitive transformer in the fashion:



$$\Bigg\} \equiv [\,-\,\underset{\substack{\cdot\\ I_1}}{1}\,-\,TF\,-\,\underset{\substack{\cdot\\ I_2}}{1}\,-\,]$$

$$\text{or } [\,-\,\underset{\substack{\cdot\\ I'}}{1}\,-\,TF\,-\,] \text{ or } [\,-\,TF\,-\,\underset{\substack{\cdot\\ I''}}{1}\,-\,]$$

In this case, the inertial impedances can be appended to both sides of an inertialess transformer or reflected entirely to either side.

This last example, then, gives us a prototype reticulation for handling the general case of a lossy, reactive transformer, namely:

$$— \underset{Z_1}{\underset{\cdot}{1}} — \mathrm{TF} — \underset{Z_2}{\underset{\cdot}{1}} —$$

where the element, TF, can be considered static and lossless.

This is usually considered as an ideal transformer for which

$$d\mathbb{E}/dt \equiv 0 \quad ; \quad \mathcal{P}_d \equiv 0$$

$$\therefore \quad \boxed{\mathbb{P}_1 \equiv e_1 f_1 \quad \equiv \quad e_2 f_2 \equiv \mathbb{P}_2}$$

Ideal Linear Transformers

We may then obtain the relations between efforts and flows for an ideal TF by considering a two-port matrix whose $\mathbb{A}$, $\mathbb{B}$, $\mathbb{C}$, $\mathbb{D}$, operat are all constant at the values a, b, c, d, respectively. Then we have:

$$\left[\begin{array}{c} e_1 \\ \hdashline f_1 \end{array}\right] = \left[\begin{array}{c:c} a & b \\ \hdashline c & d \end{array}\right] \cdot \left[\begin{array}{c} e_2 \\ \hdashline f_2 \end{array}\right]$$

The input power may be expressed:

$$\begin{aligned} e_1 f_1 &= (ae_2 + bf_2)(ce_2 + df_2) \\ &= (ae)e_2^2 + (ad + bc)\, e_2 f_2 + (bd) f_2^2 \end{aligned}$$

Since the power balance requires that $\mathbb{P}_1 \equiv \mathbb{P}_2$, then the following three conditions must hold:

$$ac \equiv 0 \quad ; \quad ad + bc \equiv 1 \quad ; \quad bd \equiv 0$$

Only two possibilities exist: the first corresponds to the normal ideal transformer; the second case to the gyrating transformer or gyrator first discussed by TELLEGEN in 1948. We can treat these two devices in parallel fashion as follows:

NORMAL TRANSFORMER
(Ideal Transformer)

Conditions:

$b \equiv c \equiv 0$; $d \equiv 1/a \equiv a$

Matrix:

$$\left[\begin{array}{c|c} a & 0 \\ \hline 0 & a' \end{array}\right]$$

Relations:

$$e_1 = ae_2$$
$$f_1 = a'f_2$$

or

$$e_1 = ae_2$$
$$f_2 = af_1$$

Determinant: $\Delta = +1$

GYRATING TRANSFORMER
(Ideal Gyrator)

Conditions:

$a \equiv d \equiv 0$; $c \equiv 1/b \equiv b'$

Matrix:

$$\left[\begin{array}{c|c} 0 & b \\ \hline b' & 0 \end{array}\right]$$

Relations:

$$e_1 = bf_2$$
$$f_1 = b'e_2$$

or

$$e_1 = bf_2$$
$$e_2 = bf_1$$

Determinant: $\Delta = -1$

The relations between these two types of ideal transformers are readily conceived in terms of a primitive <u>unit gyrator</u> or gyrating matrix, , playing a role analogous to the identity matrix, , defined as follows:

[———————] ; [—— GY ——]

$$\mathbb{M} = \mathbb{I} = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & 1 \end{array}\right] ; \quad \mathbb{M} = \mathbb{G} = \left[\begin{array}{c|c} 0 & 1 \\ \hline 1 & 0 \end{array}\right]$$

Thus $\mathbb{I}$ or [——] represents the normal or "undisturbed" bond while $\mathbb{G}$ or [—— GY ——] represents a "gyrated" bond. Moreover det $\mathbb{I} = +1$ while det $\mathbb{G} = -1$.

In these terms, the two transformers are related as follows:

—— NORMAL TRANSFORMER ——

—— TF —— O ————

$$\left[\begin{array}{c|c} a & 0 \\ \hline 0 & a' \end{array}\right] \cdot \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & 1 \end{array}\right]$$

—— GYRATING TRANSFORMER ——

—— TF —— O —— GY ——

$$\left[\begin{array}{c|c} a & 0 \\ \hline 0 & a' \end{array}\right] \cdot \left[\begin{array}{c|c} 0 & 1 \\ \hline 1 & 0 \end{array}\right]$$

This permits us to include the gyrator within the framework of ordinary ideal transformers in all that follows.

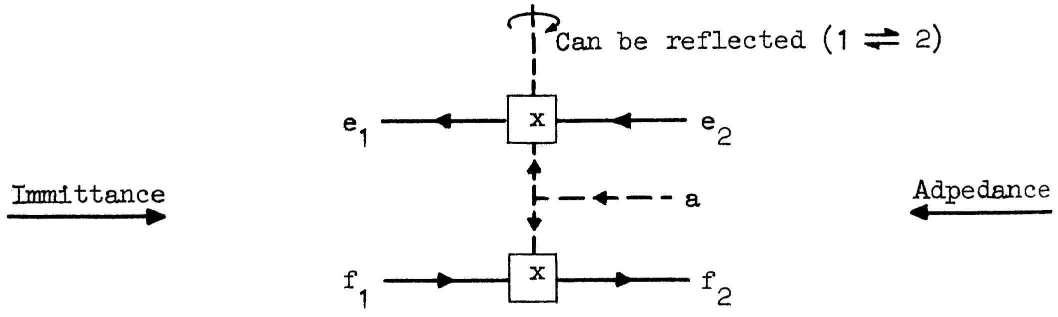
The constant a measures the (flow) transformation ratio. For the various common particularizations this is interpreted as:

MECHANICAL LEVER:	a = lever ratio
GEAR TRAIN:	a = gear ratio
HYDRAULIC JACK:	a = area ratio
ELECTRICAL TRANSFORMER:	a = turns ratio

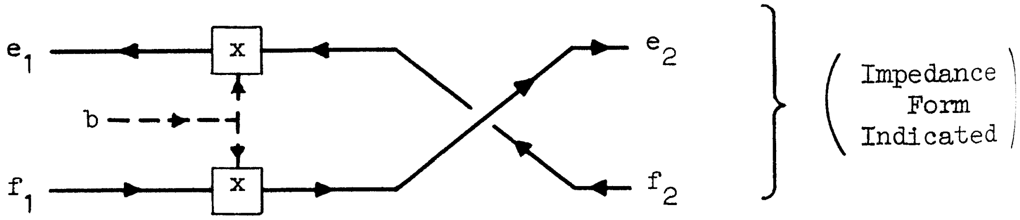
Causality of Ideal Transformers

The two causal bond diagrams for an ideal transformer would have the form:

which can both be expressed by the diagram

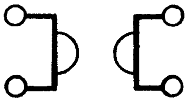


The gyrator is included within this scheme simply by crossing ove or gyrating inputs and outputs as follows:



Physical Gyrators

The gyrator has assumed great practical importance in electrical engineering, particularly in the domain of microwaves. It has been endowed with status as a standard circuit element having the symbol:



Although realizations of the purely electrical gyrator have been proposed using active elements such as amplifiers, it must be recalled that the ideal gyrator is inherently a passive device, namely one which neither stores nor dissipates energy.

From the standpoint of internal energy reticulation, the ideal electric gyrator instantaneously converts electric to magnetic energy and vice-versa. This indicates at once its enormous practical importance since a capacitor connected to one port looks like an inductor at the other port. We may see this immediately using the unit gyrator GY in the form:

$$\begin{bmatrix} e_1 \\ \hdashline f_1 \end{bmatrix} = \overbrace{\left[\begin{array}{c:c} 0 & 1 \\ \hdashline 1 & 0 \end{array}\right]}^{GY} \overset{C}{\begin{bmatrix} e_2 \\ \hdashline CDe_2 \end{bmatrix}}$$

thus

$$e_1 = CD\,e_2$$

$$f_1 = e_2$$

or

$$\boxed{e_1 = CD\,f_1}$$

whence the capacitance acts as if it were an <u>inertance</u>. The importance of this result for UHF communications cannot be overestimated!

This instantaneous conversion $\mathbb{E}_e \rightleftharpoons \mathbb{E}_m$ suggests an electrical realization in the form:

$$\frac{E_1}{I_1}\ \begin{matrix}\text{Electrostatic}\\ \text{to Mechanical}\\ \text{Transducer}\end{matrix}\ \frac{F}{V}\ \begin{matrix}\text{Mechanical to}\\ \text{Electromagnetic}\\ \text{Transducer}\end{matrix}\ \frac{E_2}{I_2}$$

since, if the two transducers were themselves ideal, the conversion would be instantaneous.

More recently, the gyrator has been approximated in solid state devices, using, for example, (a) the Hall effect, and (b) the Faraday rotation of a ferrite.

These practical realizations then permit gyration in another medi by the simple scheme:

Medium X						Medium X
$\frac{e_1}{f_1}$	X - Electric Transducer	$\frac{E_a}{I_a}$	Electrical Gyrator	$\frac{E_b}{I_b}$	Electric - X Transducer	$\frac{e_2}{f_2}$

As long as the two transducers are themselves similar, near-ideal devices, effective gyration will take place between ports 1 and 2.

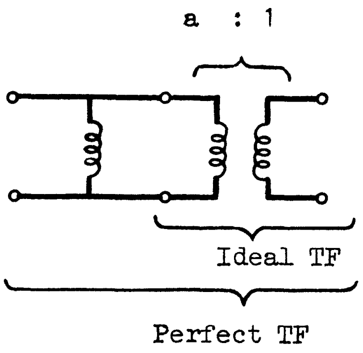
In all such practical embodiments the self-impedances Z_{11} and Z_{22} will not be zero nor the transfer-impedances Z_{12} and Z_{21} be perfectly skew-symmetric, as is the case with the ideal gyrator.

The "Perfect" Electrical Transformer

In the problem material considerable attention is given to the modeling of electrical transformers as two-port elements, It is particularly interesting to note that a so-called "perfect" transformer (with unity coupling) may be represented in the form:

$$\text{——}\ \underset{\substack{\bullet \\ \text{I}}}{0}\ \text{————}\ \text{TF}\ \text{——}$$

which in conventional electrical symbols becomes

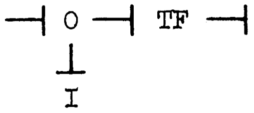


The corresponding 2-port matrix would become:

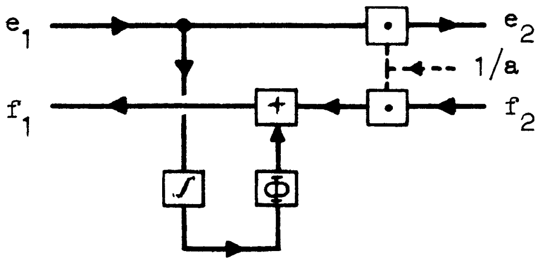
$$\mathbb{M} = \left[\begin{array}{c:c} 1 & 0 \\ \hdashline 1/LD & 1 \end{array}\right] \cdot \left[\begin{array}{c:c} a & 0 \\ \hdashline 0 & 1/a \end{array}\right]$$

$$= \left[\begin{array}{c:c} a & 0 \\ \hdashline a/LD & 1/a \end{array}\right]$$

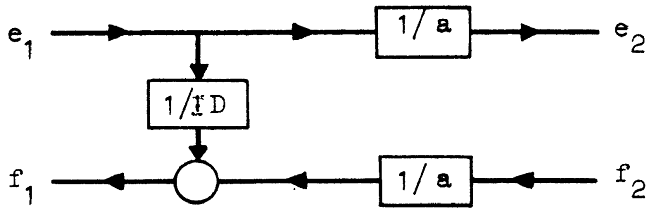
Furthermore, a possible causal representation of this structure might have the form:



with the corresponding computer flow graph



If the transformer were linear as implied, $\rightarrow \boxed{\Phi} \rightarrow$ transforms to $\underset{1/L}{\rightarrow \boxed{\cdot} \rightarrow}$. This same structure may then also be indicated in linear block diagram form as:

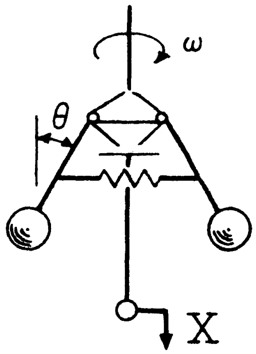


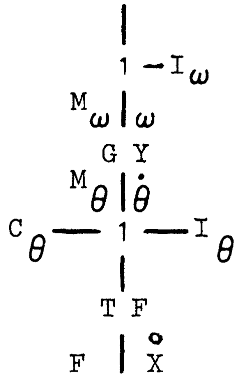
Finally, it is most important to realize that by no means is it possible to reticulate this system into a purely [E, F, 0, 1, R, C, I] structure; such a transformer is an essential two-port device.

Geometrical Constraints as Space Transformers

The introduction of the ideal transformer as a fundamental 2-por element now allows us to represent many two-and three-dimensional geometrical constraints as simple transformer couplings between the several axes involved. Some examples can serve to elucidate this possibility.

For example, the ordinary flyweight mechanism, originally introduced by James WATT and still employed as a speed sensing element, has the following form:





The important "gyroscopic transformer action" or "rotary-angular transduction" involved can be expressed by the well established dynamical relations:

$$\boxed{\begin{array}{ll} \text{CENTRIFUGAL TORQUE:} & M_\theta = [I_g(\theta)\cdot\omega]\ \omega = \Psi\cdot\omega \\ \text{GYROSCOPIC TORQUE:} & M_\omega = [I_g(\theta)\cdot\omega]\ \dot\theta = \Psi\cdot\dot\theta \end{array}}$$

However, it should be noted immediately that the existence of one of these relations immediately implies the existence of the second. Moreover the modulus Ψ of any space transformer or gyrator will always be a functional of (at least) the adjacent bond flows ($\mathscr{f}$). In a holonomic system the functional gives way to an ordinary (static) function of the coordinates or displacements ($\mathfrak{q}$), while for the rheonomic systems the functional becomes a function of the rates or flows ($\mathscr{f}$). In the example above we have a mixed situation with the modulus

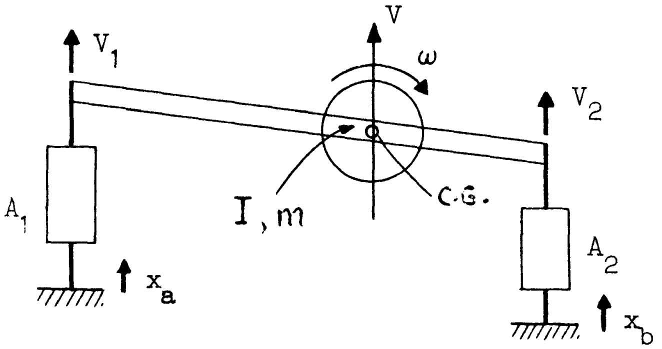
$$\Psi \equiv \Phi(\theta, \omega) = \Phi(\mathfrak{q}, \mathscr{f})$$

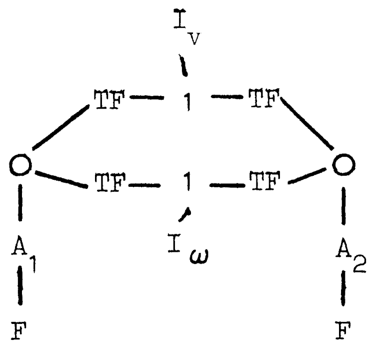
It is also interesting to note that for small changes about the equilibrium

$$\Phi = \text{const.} = b$$

and we have the previous case of an ideal gyrator. This implies that the parts of the system each side of the gyrator are mutually dual; thus for example it is C_θ and not I_θ which adds to the I_ω to give the effective inertia when looking into the flyweights from the rotary bond.

Another enlightening example is the case of a suspended vehicle or platform for which the rotary as well as linear energies must be taken into account. Such a system might have the abstract form:





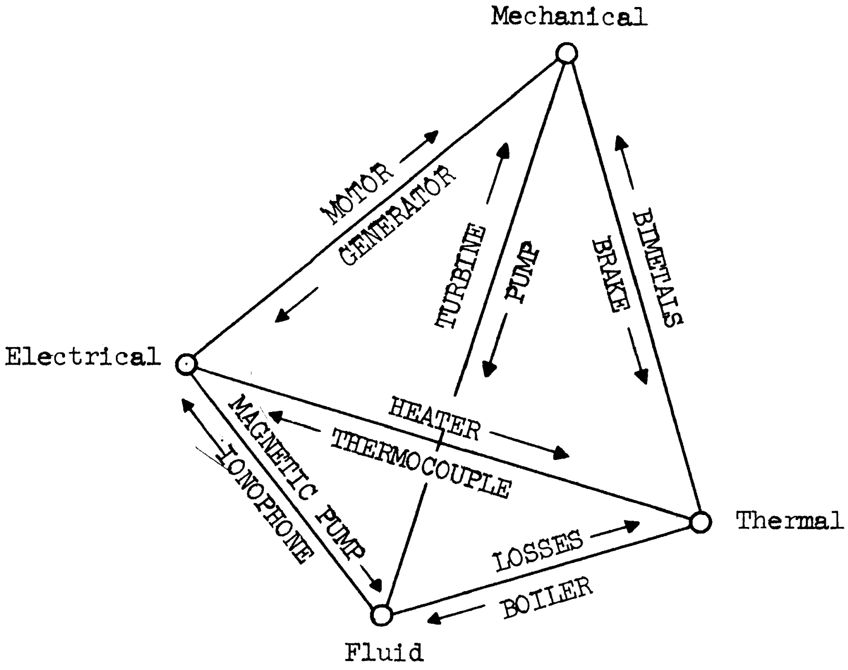
There is now no need to "transfer" or "reflect" the linear and rotary inertias to the two support points; the four transformers indicated take completely into account both the efforts and the flows associated with the two inertances.

Similar considerations will apply for all dynamical problems in two-and three-dimensional space. By these means, physical space itself becomes either a variable or a parameter, depending respectively upon whether an energy is, or is not, associated with a corresponding spatial variation. Moreover, the relation between system geometry and system topology is thereby made directly evident.

The technique of using ideal transformers to relate space axes derives from original work of Vannever BUSH, Gabriel KRON and many others. It forms the basis of highly successive use of passive electrical analogs by G. D. McCANN and his followers.

C. Energy Transduction Elements [—— TD ——]

An energy transducer or energy converter is used to convert available energy in one medium into available (or possibly unavailable) energy in another. Some of the more common forms of transducer are manifested on the transduction tetrahedron below.



TRANSDUCTION TETRAHEDRON

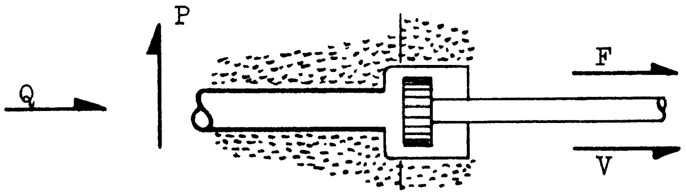
In recent times, some of the branches of this tetrahedron have even become well-established domains of engineering science such as thermoelectricity and magnetohydrodynamics (MHD). The most general relationships for such transducers are clearly no more than that for a general two-port, but considerable insight is gained by considering some useful transduction models and representations.

Below we indicate models of the form:

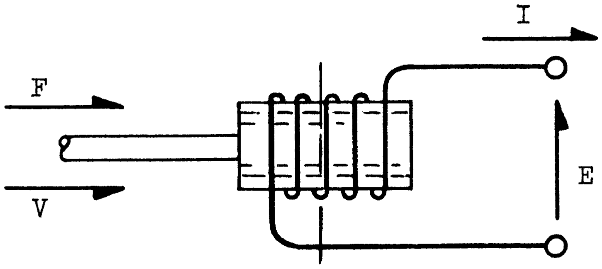
[— TM — TF — TM —]

[———— TD ————]

since a 100 per cent efficient transducer is always equivalent to an ideal transformer.



FLUID - MECHANICAL



MECHANO - ELECTRICAL

This we may see for the fluid-mechanical piston transducer since

$$F(t) = A \cdot P(t)$$

$$Q(t) = A \cdot V(t)$$

Thus:

$$\begin{bmatrix} F \\ \hdashline V \end{bmatrix} = \left[\begin{array}{c:c} A & 0 \\ \hdashline 0 & 1/A \end{array}\right] \cdot \begin{bmatrix} P \\ \hdashline Q \end{bmatrix}$$

Similarly for a solenoid transducer:

$$F(t) = (Bl) \cdot I(t)$$

$$E(t) = (Bl) \cdot V(t)$$

$$\begin{bmatrix} E \\ \hdashline I \end{bmatrix} = \left[\begin{array}{c:c} 0 & Bl \\ \hdashline 1/Bl & 0 \end{array}\right] \cdot \begin{bmatrix} F \\ \hdashline V \end{bmatrix}$$

Note in the latter case that transduction is in the form of a gyrator. In all such cases the coupling modulus serves merely as a transformation ratio, with no loss of energy. All dynamic and dissipative actions are then included in external generally nonlinear impedances, as we have noted has been standard procedure for many years in electrical transformer practice.

This gyrating model of electro-mechanical transduction is considerably more general than would appear at first glance. For example, a very refined representation of a 3-phase synchronous machine can be obtained from the directly generalized relations

$$\boxed{\begin{aligned} M &= \boldsymbol{\Psi} \cdot \mathbb{I} \\ \mathbb{E} &= \boldsymbol{\Psi} \cdot N \end{aligned}}$$

where $\mathbb{E} = [E_1(t), E_2(t), E_3(t)]$, $\mathbb{I} = [I_1(t), I_2(t), I_3(t)]$ are the instantaneous phase voltages and currents of the three phases and the modul

$$\boldsymbol{\Psi} = \boldsymbol{\Psi}(\mathbb{I}, N)$$

has the three phase components

$$\begin{aligned} \boldsymbol{\Psi}_1 &= \Phi_1(\mathbb{I}) \cdot \sin(kNt - 0\pi) \\ \boldsymbol{\Psi}_2 &= \Phi_2(\mathbb{I}) \cdot \sin(kNt - \tfrac{2}{3}\pi) \\ \boldsymbol{\Psi}_3 &= \Phi_3(\mathbb{I}) \cdot \sin(kNt - \tfrac{4}{3}\pi) \end{aligned}$$

Yet another generalization is possible following standard acousti practices, particularly where the transducer is approximately linear, namel an impedance matrix representation, which in this application is originally due to Henri POINCARE, and is conventionally written:

$$\left[\begin{array}{c} e_1 \\ \hline e_2 \end{array}\right] = \left[\begin{array}{c:c} \mathbb{Z}_1 & \mathbb{T}_{12} \\ \hdashline \mathbb{T}_{21} & \mathbb{Z}_2 \end{array}\right] \cdot \left[\begin{array}{c} f_1 \\ \hline f_2 \end{array}\right]$$

$\mathbb{Z}_1$, $\mathbb{Z}_2$ are the (self) impedances of Medium I and Medium II, respectivel while $\mathbb{T}_{12}$ and $\mathbb{T}_{21}$ are the so-called transduction operators that describ the coupling between Medium I and Medium II. The dramatic history behind such developments is discussed delightfully by HUNT; some recent applications to electrical machinery are treated by RIDEOUT and SWIFT.

Another form of description has evolved out of the field of fluid machinery in the form of relations based upon dynamic similarity, which are used extensively for problems involving pumps, turbines, and aircraft and marine propellers.

The general turbomachine may be depicted in the causal form:

$$\Big| \frac{M}{\omega} \text{ Turbomachine } \frac{P}{Q} \Big|$$

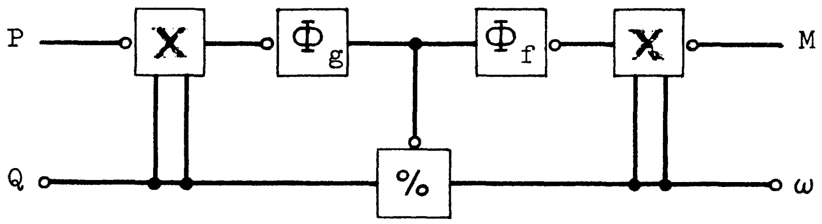
Here it is only necessary to realize that a geometrically similar flow field will exist for any operating point along a line of similitude defined by:

$$\boxed{Q = r\omega}$$

This constraint imposes severe restrictions upon the possible form of the resultant machine characteristics, giving as one of several possibilities the pair of relations

$$\boxed{\begin{array}{l} M = f(r) \cdot \omega^2 \\ P = g(r) \cdot Q^2 \end{array}}$$

The corresponding computing structure may be realized in the form:



Particularly simple models result upon taking

$$f(r) \cong a + br + cr^2 + \cdots$$
$$g(r) \cong (e/r^2) + (f/r) + h + \cdots$$

where the series coefficients (a, b, c, ..., e, f, h, ...) depend upon the geometry of the device, including the possibly variable blading and other modulating parameters.

Lastly, we may obtain certain general results for the case of constant efficiency reciprocal transducers operating over a small range. A per-unit notation reveals:

$$\frac{d\mathbb{P}_1}{\mathbb{P}_{01}} = u_1 + v_1 \quad ; \quad \frac{d\mathbb{P}_2}{\mathbb{P}_{02}} = u_2 + v_2$$

where $u = de/e = d(\ln e)$ and $v = df/f = d(\ln f)$

thus

$$\left[\begin{array}{c} u_1 \\ \hline v_1 \end{array}\right] = \left[\begin{array}{c|c} a & b \\ \hline c & d \end{array}\right] \cdot \left[\begin{array}{c} u_2 \\ \hline v_2 \end{array}\right]$$

If efficiency is constant:

$$a + c \equiv 1 \quad ; \quad b + d \equiv 1$$

and if the transducer is reciprocal:

$$ad - bc \equiv \pm 1$$

These three conditions reduce the matrix to the two forms

$$\left[\begin{array}{c|c} a & a - 1 \\ \hline 1 - a & 2 - a \end{array}\right] \quad \text{or} \quad \left[\begin{array}{c|c} b - 1 & b \\ \hline 2 - b & 1 - b \end{array}\right]$$

A value $a \equiv 1$ or $b \equiv 1$ then yields 100 per cent efficiency, and corresponds to the ideal transformer and gyrator, respectively.

Besides their use as power-level energy converters, transducers form the essential elements of most continuous measuring instruments for physical and engineering processes. A very complete list of instrument transducers has been prepared by D. B. KRET of the Du Mont Laboratories, entitled Transduc A Compilation Useful Primarily in Oscillography.

kground Reading -- Electro-Acoustic Transducers

1) HUNT, F. V.: Electroacoustics: The Analysis of Transduction and Its Historical Background (1954)

A superb account of the history of the subject.

2) FISCHER, F. A.: Fundamentals of Electroacoustics (1955)

Presents the physical principles underlying electro-acoustic transduction.

ckground Reading -- Electro-Mechanical Transducers

1) RIDEOUT, V. C.: Analysis of D. C. Rotating Machines as Two-Port Networks, AIEE Conference Paper, CP-57-760 (1957).

2) SWIFT, W. B.: Analysis of D. C. Rotating Machines as Two-Port Networks, AIEE Conference Paper, CP-57-761 (1957)

These companion papers give an excellent introduction to the d. c. machine as a transducer.

3) WHITE, D. C. and WOODSON, H. H.: Electromechanical Energy Conversion (1959)

The current definitive work in this subject.

XVI. Energy Transmission Elements

A. The Two-Port Element: [—— TM ——]

To convey or transmit energy from one place where it is available -- the source -- to another place where it is to be used -- the load -- a 2-port transmitting medium or transmission element [— TM —] is required. Some common realizations might be indicated as in the following table:

FORM OF ENERGY	TRANSMISSION ELEMENT
Fluid	Pipe or Duct
Mechanical	Rod or Shaft
Electrical	Wire or Conductor

The absolutely ideal •TM• is the simple power bond:

[————]

But all real transmission elements possess static and dynamic properties resulting in the dissipation, scattering, and storage of energy.

Thus, in the case of power transmission, nonideal behavior manifests itself in power losses; while in devices for the transmission of signals and information, all real transmission links will delay, distort, attenuate, scatter, and contaminate the desired signals.

The paragraphs below deal with, and distinguish between both situations, and relate behavior to the limiting cases of pure wavelike and pure diffusive transmission.

Steady Losses in Power Transmission

Within and around any •TM• element, part of the available energy bei transported is continuously consumed and converted into heat. Under steady operating conditions, along the entire transmission path, from the source to the load, there will be a net convergence of the Poynting vector, $\overline{\mathbb{P}}$, resulting in a power gradient approximately parallel to the transmitter and a continuous decrease in power level along its length.

No practical devices exist which can transmit power over space ithout such corresponding parasitic guidance- or support-losses. However he effects of these power losses on the overall flow of power in engineerng systems are usually restricted within fairly narrow limits.

On the one hand, since any resistance in the ·TM· consumes useful ower and thus lowers the efficiency of transmission, it is generally unconomical to permit too high a value of transmission resistance. On the ther hand, since lower resistance usually implies larger quantities of transission materials, there are also lower practical limits to power loss.

As a result of the above considerations and other factors, the teady rated percent power loss of most transmitting elements will be modest. n any case, all such losses in energy can be reduced to combinations of the wo forms:

Loss in Effort : RESISTANCE : $[\,-\underset{\substack{\bullet\\ R}}{1}-\,]$

Loss in Flow : LEAKAGE : $[\,-\underset{\substack{\bullet\\ G}}{0}-\,]$

The resultant steady loss of any transmission system must therefore e capable of representation in the form:

$$[\,-\underset{\substack{\bullet\\ R}}{1}-\underset{\substack{\bullet\\ G}}{0}-\,]^n$$

here the index n is taken sufficiently large. Under most circumstances, by ombining resistance and conductance relations, the overall loss relations y be modeled:

$$[\,-\underset{\substack{\bullet\\ R_e}}{1}-\underset{\substack{\bullet\\ G_e}}{0}-\,]$$

here R_e and G_e are equivalent resistances and conductances.

namic Effects in Power Transmission

Under transient conditions of operation, the field effects of energy torage, in the form of inertance and capacitance distributed along the transssion path, will produce significant effects upon system behavior. These ill be manifested in local variations in the distribution of power and energy ver the extent of the transmitter.

While a transmission system can be roughly characterized by a reticular model of low order, whenever the wavelengths of the power state variables [e(t), f(t)] become small by comparison to the length of the transmission path, the continuous nature of the •TM• element must be taken into account.

The Concept of Reticular Loss Junctions

An actual length of transmission line in any medium may be conceived as having all effort and flow losses concentrated in reticular fashion at the upstream and/or downstream ends in the following manner:

$$\overbrace{\left[\text{— Loss Element — Lossless Transmitter — Loss Element —}\right]^{n}}^{[\text{——— Actual Transmitter with Losses ———}]}$$

Clearly, if this situation is assumed to hold for a sufficiently large number, n, of appropriately small transmitter segments, any real transmission system may be approximated to an arbitrary degree of practical accuracy. Frequently only a few such elements are necessary. The particular sections or junctions at which all energy losses are presumed to occur may be called loss junctions.

It is particularly important to realize that while certain forms of linear loss may sometimes be handled in other analytical fashions, this use of loss junctions becomes mandatory for the representation of essential non-linear and/or reticular resistance and leakage (NONLINEAR: e.g.: electrical: corona loss; fluid: turbulent loss; thermal: radiation loss) (RETICULAR: e.g electrical: suspension insulators; fluid: bend losses; mechanical: bearing resistance).

Under rapid transient conditions, many factors conspire to make actual dynamic losses somewhat more complex. The flow will usually vary from point to point due to distributed capacitance; this will necessarily cause the local losses to vary. If the flow has sufficiently high frequency components, additional resistance and scattering phenomena are present which do not manifest themselves at lower frequencies or in the steady state. Indeed, all experimental evidence in every field medium indicates that the higher frequency components attenuate more rapidly than lower frequency signals.

B. The Canonical Transmitter Matrix ($\mathbb{T}$)

Chains of Symmetric 2•Ports: Reticular Transmitters

Consider the general asymmetric 2•port:

$$[\text{——} \ \mathrm{TP} \ \text{——}]$$

A symmetric 2•port can always be produced by connecting two such two-ports back-to-back in either of the alternative fashions:

$$\text{——} \underset{\rightarrow}{\mathrm{TP}} \text{——} \underset{\leftarrow}{\mathrm{PT}} \text{——} \qquad \text{or} \qquad \text{——} \underset{\leftarrow}{\mathrm{PT}} \text{——} \underset{\rightarrow}{\mathrm{TP}} \text{——}$$

But we know that for any linear reciprocal, symmetric 2•port only <u>two</u> of the four operators can be taken as independent.

A canonical form for the transmission matrix of all such <u>reversible</u> elements can be obtained through two new <u>defined</u> operators:

Propagation Operator $\Gamma \equiv \cosh^{-1} \mathbb{A}$

Characteristic Impedance $\mathbb{Z}_o \equiv \sqrt{\mathbb{B} / \mathbb{C}} \equiv 1/\mathbb{Y}_o$

which results in a final transmission matrix:

$$\mathbb{M} = \mathbb{T} \equiv \left[\begin{array}{c:c} \cosh \Gamma & \mathbb{Z}_o \sinh \Gamma \\ \hdashline \mathbb{Y}_o \sinh \Gamma & \cosh \Gamma \end{array}\right]$$

Consider now a chain or cascade of <u>n</u> such identical reversible elements;

$$[\text{——} \underset{1}{\mathrm{TT}} \text{——} \underset{2}{\mathrm{TT}} \text{——} \quad \cdots \quad \text{——} \underset{n}{\mathrm{TT}} \text{——}] \equiv [\text{——} \mathrm{TT} \text{——}]^n$$

We may consider any such system as a <u>reticular</u> <u>transmitter</u>.

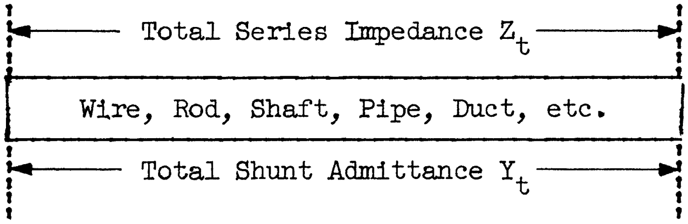
Then:

$$\mathbb{M}_n = \mathbb{T}^n = \left[\begin{array}{c:c} \cosh \Gamma & \mathbb{Z}_o \sinh \Gamma \\ \hdashline \mathbb{Y}_o \sinh \Gamma & \cosh \Gamma \end{array}\right]^n = \left[\begin{array}{c:c} \cosh n\Gamma & \mathbb{Z}_o \sinh n\Gamma \\ \hdashline \mathbb{Y}_o \sinh n\Gamma & \cosh n\Gamma \end{array}\right]$$

The value of the $\mathbb{T}$ matrix is now obvious since it is the only form for $\mathbb{M}$ which preserves its nature upon multiplication; these properties result directly from the identities of hyperbolic trigonometry.

Continuous Uniform Transmitters

Consider a uniform one-dimensional linear transmitter:



Since the overall structure is symmetric, this system has a $\mathbb{T}$ -matrix:

$$\mathbb{T} = \left[\begin{array}{c:c} \cosh \boldsymbol{\Gamma} & \mathbb{Z}_o \sinh \boldsymbol{\Gamma} \\ \hdashline \mathbb{Y}_o \sinh \boldsymbol{\Gamma} & \cosh \boldsymbol{\Gamma} \end{array}\right]$$

Moreover, we may consider this system to be reticulated into n identical symmetric micro-elements to obtain

$$\mathbb{T} = \left[\begin{array}{c:c} \cosh \frac{1}{n} \boldsymbol{\Gamma} & \mathbb{Z}_o \sinh \frac{1}{n} \boldsymbol{\Gamma} \\ \hdashline \mathbb{Y}_o \sinh \frac{1}{n} \boldsymbol{\Gamma} & \cosh \frac{1}{n} \boldsymbol{\Gamma} \end{array}\right]^n$$

If n is sufficiently large the structure of the micro-element is not critica
Either of the following two molecules would serve as micro-elements:

$$[\text{ ——— TEE ——— }]^n$$

$$[\text{ —— }\underset{\bullet}{1}\text{ —— }\underset{\bullet}{0}\text{ —— }\underset{\bullet}{1}\text{ —— }]^n$$

$$Z_t/2n \quad Y_t/n \quad Z_t/2n$$

$$\left[\begin{array}{c|c} 1+\frac{Z_t\,Y_t}{2n^2} & \frac{Z_t}{n}+\frac{Z_t^2\,Y_t}{4n^3} \\ \hline Y_t/n & 1+\frac{Z_t\,Y_t}{2n^2} \end{array}\right]^n$$

$$[\text{ ——— PI ——— }]^n.$$

$$[\text{ —— }\underset{\bullet}{0}\text{ —— }\underset{\bullet}{1}\text{ —— }\underset{\bullet}{0}\text{ —— }]^n$$

$$Y_t/2n \quad Z_t/n \quad Y_t/2n$$

$$\left[\begin{array}{c|c} 1+\frac{Z_t\,Y_t}{2n^2} & Z_t/n \\ \hline \frac{Y_t}{n}+\frac{Z_t\,Y_t^2}{4n^3} & 1+\frac{Z_t\,Y_t}{2n^2} \end{array}\right]^n$$

For both cases, there follows directly, for $\cosh(\frac{1}{n}\Gamma) \equiv \mathbb{A}$ and $\mathbb{Z}_o^2 \equiv \mathbb{B}/\mathbb{C}$, the values:

$$\cosh(\tfrac{1}{n}\Gamma) = 1 + (\tfrac{1}{2n^2})Z_tY_t \quad ;$$

$$\mathbb{Z}_o^2 = (Z_t/Y_t)[1 + (\tfrac{1}{4n^2})Z_tT_t]$$

$$\cosh(\tfrac{1}{n}\Gamma) = 1 + (\tfrac{1}{2n^2})Z_tY_t$$

$$\mathbb{Z}_o^2 = (Z_t/Y_t)/[1 + (\tfrac{1}{4n^2})Z_tY_t]$$

We may determine the $\mathbb{T}$ matrix for the continuous transmitter merely by letting the number of microelements, n, become infinite. In this case, for either •TEE• or •PI• elements there results:

$$\lim_{n \to \infty} \Gamma^2 = Z_t \cdot Y_t$$

$$\lim_{n \to \infty} \mathbb{Z}^2 = Z_t / Y_t$$

yielding the final results

Overall Propagation Operator $\Gamma = \sqrt{Z_t \cdot Y_t}$

Characteristic Impedance $\mathbb{Z}_o = \sqrt{Z_t / Y_t}$

Substituting these results into the system $\mathbb{T}$-matrix gives:

$$\mathbb{T} = \left[\begin{array}{c:c} \cosh\sqrt{Z_t \cdot Y_t} & \sqrt{Z_t/Y_t}\,\sinh\sqrt{Z_t \cdot Y_t} \\ \hdashline \sqrt{Y_t/Z_t}\,\sinh\sqrt{Z_t \cdot Y_t} & \cosh\sqrt{Z_t \cdot Y_t} \end{array}\right]$$

Note that specification of either of two pairs of operators, (Z_t, Y_t) or $(\boldsymbol{\Gamma}, \mathbb{Z}_o)$ is sufficient to specify $\mathbb{T}$ in any instance. It is also obviou that both direct and converse relations exist between these two sets in the fashion:

$$\boldsymbol{\Gamma} = \sqrt{Z_t \cdot Y_t} \quad ; \quad Z_t = \boldsymbol{\Gamma} \cdot \mathbb{Z}_o$$

$$\mathbb{Z}_o = \sqrt{Z_t / Y_t} \quad ; \quad Y_t = \boldsymbol{\Gamma} / \mathbb{Z}_o$$

The $\mathbb{T}$-matrix above is the canonical form for all linear uniform transmission elements, regardless of the nature of Z_t and Y_t. We shall find it convenient in the treatment below to deal with two limiting cases of transmitters, namely:

—— WAVELIKE —— TM	—— DIFFUSIVE —— TM
$Z_t \equiv I_t D$	$Z_t \equiv R_t$
$Y_t \equiv C_t D$	$Y_t \equiv C_t D$

C. Generalized Transmitters and Wavelike Transmitters

Consider the one-dimensional generalized transmission process governed by the matrix differential equation:

$$\boxed{(d\,\mathbb{S}/dx) + \mathbb{N}\,\mathbb{S} = 0}$$

where $\mathbb{S} \equiv \left[\begin{array}{c} e \\ \hline f \end{array}\right]$ is the power state vector. This relation is simply a formulation of the intuitive relation:

$$\mathbb{S} + d\,\mathbb{S} = [\,\mathbb{I} + \mathbb{N}\;dx]\;\mathbb{S}$$

Let us assume a matrix solution of the form:

$$\mathbb{S}_x = \mathbb{M}_x\,\mathbb{S}_2$$

where $\mathbb{S}_2$ is the downstream power state and $\mathbb{M}_x$ is the 2-port transmission matrix for a length x of •TM•.

Then:

$$d\,\mathbb{S}/dx = (\,\mathbb{M}/dx) \cdot \mathbb{S}_2$$

and

$$\mathbb{N}\,\mathbb{S} = \mathbb{N}\;\mathbb{M}\;\mathbb{S}_2$$

which yield an equivalent differential equation for $\mathbb{M}$, namely

$$(d\,\mathbb{M}/dx) + \mathbb{N}\,\mathbb{M} = 0$$

This relation could be used directly to determine differential equations for the $\mathbb{A}$, $\mathbb{B}$, $\mathbb{C}$, $\mathbb{D}$ elements, but a much more instructive procedure, at least for wavelike transmitters, is to first transform the power state $\mathbb{S}$ to a characteristic state $\mathbb{R}$ in terms of characteristic (or scattering) variables (u,v) where:

$$\begin{array}{ccccccc} \mathbb{R} & = & \mathbb{H} & \cdot & \mathbb{E} & \cdot & \mathbb{S} \\ \left[\begin{array}{c} u \\ \hline v \end{array}\right] & = & \frac{1}{\sqrt{2}}\left[\begin{array}{c|c} 1 & 1 \\ \hline -1 & 1 \end{array}\right] & \cdot & \left[\begin{array}{c|c} G_o & 0 \\ \hline 0 & \sqrt{R_o} \end{array}\right] & \cdot & \left[\begin{array}{c} e \\ \hline f \end{array}\right] \end{array}$$

with the converse relations

$$\mathcal{S} = \mathbb{E}^{*} \cdot \mathbb{H} \cdot \mathbb{R}$$

$$\left[\begin{array}{c} e \\ \hline f \end{array}\right] = \left[\begin{array}{c:c} \sqrt{R_o} & 0 \\ \hline 0 & \sqrt{G_o} \end{array}\right] \frac{1}{\sqrt{2}} \left[\begin{array}{c:c} 1 & 1 \\ \hline 1 & -1 \end{array}\right] \left[\begin{array}{c} u \\ \hline v \end{array}\right]$$

We can relate the characteristic state, $\mathbb{R}_x$, at any upstream point, x, to the downstream state, $\mathbb{R}_2$, by the expression

$$\mathbb{R}_x = \underbrace{\mathbb{H}\,\mathbb{E}\,\mathbb{M}\,\mathbb{E}^{*}\mathbb{H}}\,\mathbb{R}_2$$

$$\mathbb{R}_x = \mathbb{Q} \cdot \mathbb{R}_2$$

Thus the characteristic matrix, $\mathbb{Q}$, (which is directly related to the scattering matrix) plays the same role for characteristic variables as the transmission matrix $\mathbb{M}$ plays for power variables.

If we define a matrix $\mathbb{T}$ analogous to $\mathbb{N}$ above, by the relation:

$$\mathbb{T} \equiv \mathbb{H}\,\mathbb{E}\,\mathbb{N}\,\mathbb{E}^{*}\mathbb{H}$$

then the differential equation for the general characteristic matrix, $\mathbb{Q}$, is given by:

$$(d\,\mathbb{Q}/dx) + \mathbb{T}\,\mathbb{Q} = 0$$

This last equation can then be solved for the case is a uniform lossless wavelike transmitter in a very direct fashion as follows:

For the wavelike TM

$$\mathbb{N} \equiv \left[\begin{array}{c:c} 0 & \mathcal{L}D \\ \hline \mathcal{C}D & 0 \end{array}\right]$$

Then $\mathbb{T}$ can be found to be

$$\mathbb{T} \equiv \left[\begin{array}{c:c} \sqrt{\mathcal{L}\mathcal{C}}D & 0 \\ \hline 0 & \sqrt{\mathcal{L}\mathcal{C}}D \end{array}\right] = \left[\begin{array}{c:c} \gamma & 0 \\ \hline 0 & -\gamma \end{array}\right]$$

if the scaling constants are taken as $R_o \equiv 1/G_o \equiv \sqrt{\mathcal{L}/\mathcal{C}}$.

Writing:

$$\mathbb{Q} = \left[\begin{array}{c:c} \mathcal{A} & \mathcal{B} \\ \hdashline \mathcal{C} & \mathcal{D} \end{array}\right]$$

we find:

$$\mathbb{Q}^{\triangledown} = \quad d\,\mathbb{Q}\,/x \quad = \left[\begin{array}{c:c} \mathcal{A}' & \mathcal{B}' \\ \hdashline \mathcal{C}' & \mathcal{D}' \end{array}\right]$$

$$\mathbb{T}\mathbb{Q} = \left[\begin{array}{c:c} \gamma & 0 \\ \hdashline 0 & -\gamma \end{array}\right] \left[\begin{array}{c:c} \mathcal{A} & \mathcal{B} \\ \hdashline \mathcal{C} & \mathcal{D} \end{array}\right] = \left[\begin{array}{c:c} \gamma\mathcal{A} & \gamma\mathcal{B} \\ \hdashline -\gamma\,\mathcal{C} & -\gamma\mathcal{D} \end{array}\right]$$

Adding matrices we obtain:

$$\left[\begin{array}{c:c} \mathcal{A}' + \gamma\mathcal{A} & \mathcal{B}' + \gamma\mathcal{B} \\ \hdashline \mathcal{C}' - \gamma\mathcal{C} & \mathcal{D}' - \gamma\mathcal{D} \end{array}\right] = 0$$

From the primitive statement $\mathbb{Q} + d\,\mathbb{Q} = (\,\mathbb{I} + \mathbb{T}\,dx)\,\mathbb{Q}$, taking into account the nature of $\mathbb{T}$, then necessarily:

$$\mathcal{B} \equiv \mathcal{C} \equiv 0$$

For $\mathcal{A}$ and $\mathcal{D}$, we have the parallel developments:

$$\begin{array}{c:c}
\mathcal{A}' + \gamma\mathcal{A} = 0 & \mathcal{D}' - \gamma\mathcal{D} = 0 \\
\mathcal{A}' = -\gamma\mathcal{A} & \mathcal{D}' = +\gamma\mathcal{D} \\
\dfrac{\mathcal{A}'}{\mathcal{A}} = \dfrac{d}{dx}\ \ln\mathcal{A} = -\gamma & \dfrac{\mathcal{D}'}{\mathcal{D}} = \dfrac{d}{dx}\ \ln\mathcal{D} = \gamma \\
d\ln\mathcal{A} = -\gamma\,dx & d\ln\mathcal{D} = \gamma\,dx \\
\displaystyle\int_0^{\mathcal{A}} d\ln\mathcal{A} = \int_x^0 -\gamma\,dx & \displaystyle\int_0^{\mathcal{D}} d\ln\mathcal{D} = \int_x^0 \gamma\,dx \\
\ln\mathcal{A} = \gamma\,x & \ln\mathcal{D} = -\gamma\,x \\
\boxed{\mathcal{A} = e^{\gamma x} = e^{\Gamma}} & \boxed{\mathcal{D} = e^{-\gamma x} = e^{-\Gamma}}
\end{array}$$

This yields the final universal transmission relation:

$$\boxed{\mathbb{Q} = \left[\begin{array}{c|c} e^{\Gamma} & 0 \\ \hline 0 & e^{-\Gamma} \end{array}\right]}$$

which is in the form of a transformer.

This matrix is actually applicable to a broader class of transmitter than merely the wavelike variety, but the latter instance is our present concern.

For this wavelike case:

$$\Gamma = \sqrt{\mathcal{I}\mathcal{C}} \cdot x D = \sqrt{(\mathcal{I} x)(\mathcal{C} x)}\, D = \sqrt{I_t C_t}\, D = TD$$

where T is simply the <u>wave propagation time</u> from upstream to downstream port. The corresponding $\mathbb{Q}$ -matrix is

$$\boxed{\mathbb{Q} = \left[\begin{array}{c|c} e^{TD} & 0 \\ \hline 0 & e^{-TD} \end{array}\right]}$$

As a last step we can transform this result back to the transmission matrix $\mathbb{M}$, to obtain

$$\mathbb{M} = \mathbb{E}^{*}\mathbb{H}\,\mathbb{Q}\,\mathbb{H}\,\mathbb{E}$$

or

$$\boxed{\mathbb{M} = \left[\begin{array}{c|c} \cosh TD & R_o \sinh TD \\ \hline G_o \sinh TD & \cosh TD \end{array}\right]}$$

with

$$R_o \equiv 1/G_o \equiv \sqrt{I_t/C_t}\ ; \quad T \equiv \sqrt{I_t \cdot C_t}$$

D. Ideal Wavelike Transmitters

We are here concerned with continuous elements of the form:

WIRE, PIPE, ROD, SHAFT, DUCT, etc.

——— TM ———

having the fundamental transmission matrix:

$$\left[\begin{array}{c:c} \cosh TD & R_o \sinh TD \\ \hdashline G_o \sinh TD & \cosh TD \end{array}\right]$$

T = Propagation Time $= \sqrt{I_t \cdot C_t}$

R_o = $1/G_o$ = Surge Resistance $= \sqrt{I_t / C_t}$

This same matrix can be used for general operational analyses, describing system behavior in either the frequency or time domains. Contrary to common opinion, transients and vibrations in such transmitters become extremely simple to investigate so long as the most effective description is employed.

Use of Compatible or Homogeneous Variables

A significant simplification takes place if the variables are transformed in such a way that both effort and flow are measured in the same units. This normalization will then eliminate one of the parameters in the transmission matrix, as indicated in the following tabulation:

ITEM	NORMALIZATION SCHEMES		
	SCHEME A	SCHEME B	SCHEME C
Normalized State Vector $\mathbb{S} = \left[\begin{array}{c} \epsilon \\ \hdashline \phi \end{array}\right]$	$\epsilon = e$ $\phi = R_o f$	$\epsilon = G_o e$ $\phi = f$	$\epsilon = \sqrt{G_o \cdot e} = \sqrt{\mathbb{P}_e}$ $\phi = \sqrt{R_o \cdot f} = \sqrt{\mathbb{P}_f}$
Normalized Transmission Matrix $\mathbb{T}$	$= \left[\begin{array}{c:c} \cosh TD & \sinh TD \\ \hdashline \sinh TD & \cosh TD \end{array}\right]$	$= \left[\begin{array}{c:c} \cosh TD & \sinh TD \\ \hdashline \sinh TD & \cosh TD \end{array}\right]$	$= \left[\begin{array}{c:c} \cosh TD & \sinh TD \\ \hdashline \sinh TD & \cosh TD \end{array}\right]$

Intrinsic Wavelike Transmission Matrix

We may consider the reduction to intrinsic variables as a factoring of the transmission matrix in the form:

$$\mathbb{M} = \left[\begin{array}{c|c} \cosh TD & R_o \sinh TD \\ \hline G_o \sinh TD & \cosh TD \end{array}\right]$$

$$\mathbb{M} = \left[\begin{array}{c|c} \sqrt{R_o} & 0 \\ \hline 0 & \sqrt{G_o} \end{array}\right] \left[\begin{array}{c|c} \cosh TD & \sinh TD \\ \hline \sinh TD & \cosh TD \end{array}\right] \left[\begin{array}{c|c} \sqrt{G_o} & 0 \\ \hline 0 & \sqrt{R_o} \end{array}\right]$$

$$\mathbb{M} = \mathbb{K} \cdot \mathbb{T} \cdot \mathbb{K}$$

The matrix $\mathbb{K}$ is that for an ideal transformer of modulus $\sqrt{R_o}$: while the matrix $\mathbb{T}$ represents the ideal wavelike transmitter with transit time T. We shall find it convenient to adopt the symbolism:

$$\mathbb{T} = \left[\begin{array}{c|c} C & S \\ \hline S & C \end{array}\right]$$

where $C \equiv \cosh TD$ and $S \equiv \sinh TD$. Note that the well-known identity

$$C^2 - S^2 \equiv 1$$

merely expresses the reciprocity condition for such transmitters.

Characteristic or Scattering Variables

Consider the transformation:

$$\left[\begin{array}{c} u \\ \hline v \end{array}\right] = \frac{1}{\sqrt{2}} \left[\begin{array}{c|c} 1 & 1 \\ \hline 1 & -1 \end{array}\right] \cdot \left[\begin{array}{c} e \\ \hline f \end{array}\right]$$

or:

$$u = (1/\sqrt{2})\,(e + f) \qquad \text{[Sum]}$$

$$v = (1/\sqrt{2})\,(e - f) \qquad \text{[Diff.]}$$

The inverse transformation is readily determined as:

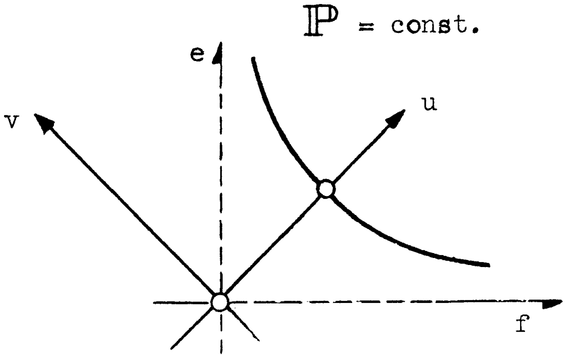
$$\left[\begin{array}{c} e \\ \hline f \end{array}\right] = \frac{1}{\sqrt{2}} \left[\begin{array}{c|c} 1 & 1 \\ \hline 1 & -1 \end{array}\right] \cdot \left[\begin{array}{c} u \\ \hline v \end{array}\right]$$

or: $e = (1/\sqrt{2})\,(u + v)$ [Sum]

$f = (1/\sqrt{2})\,(u - v)$ [Diff.]

The instantaneous power, $\mathbb{P}(t)$, can now be expressed:

$$\mathbb{P} = e \cdot f = u^2/2 - v^2/2 = \overset{\Rightarrow}{\mathbb{P}}_t - \overset{\Leftarrow}{\mathbb{P}}_r$$



Maximum $\mathbb{P}$ always occurs when $u^2 \rightarrow$ max , $v^2 \rightarrow$ min $\rightarrow 0$

Zero $\mathbb{P}$ corresponds to $v = \pm u$

Here:

TRANSMITTED POWER (or COFLUENT Power) : $\overset{\Rightarrow}{\mathbb{P}}_t \equiv u^2/2$

REFLECTED POWER (or COUNTERFLUENT Power) : $\overset{\Leftarrow}{\mathbb{P}}_r \equiv v^2/2$

Thus u(t) measures the instantaneous value of the downstream flowing or <u>cofluent</u> power and v(t), the instantaneous value of upstream flowing or <u>counterfluent</u> power.

The variables (u,v) are known in electrical science as <u>scattering parameters</u> and in fluid mechanics as <u>characteristic variables</u>. We may define a <u>characteristic state vector</u>, $\mathbb{R}$, by the definition:

$$\mathbb{R} \equiv \left[\begin{array}{c} u \\ \hline v \end{array}\right]$$

which is analogous to the normal state vector:

$$\mathbb{S} \equiv \left[\begin{array}{c} e \\ \hline f \end{array}\right]$$

Then the relation between these two measures of state is given by the converse pair of relations:

$$\begin{array}{cccc|cccc} \mathbb{R} & = & \mathbb{H} & \cdot\ \mathbb{S} & \mathbb{S} & = & \mathbb{H} & \mathbb{R} \\ \left[\begin{array}{c} u \\ \hline v \end{array}\right] & = & \frac{1}{\sqrt{2}}\left[\begin{array}{c|c} 1 & 1 \\ \hline 1 & -1 \end{array}\right] & \cdot \left[\begin{array}{c} e \\ \hline f \end{array}\right] & \left[\begin{array}{c} e \\ \hline f \end{array}\right] & = & \frac{1}{\sqrt{2}}\left[\begin{array}{c|c} 1 & 1 \\ \hline 1 & -1 \end{array}\right] & \cdot \left[\begin{array}{c} e \\ \hline v \end{array}\right] \end{array}$$

Note that since $\mathbb{H}^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \equiv \mathbb{I}$, the scattering operator, $\mathbb{H}$, is analogous to the gyrating operator $\mathbb{G}$, in being another of the (many) square roots of $\mathbb{I}$, and therefore representing a duality transformation.

<u>Characteristic Relations for Wavelike Transmission:</u>

Let us now apply the scattering matrix $\mathbb{H}$ fore and aft of an ideal transmitter $\mathbb{T}$, thus:

$$\begin{aligned} \mathbb{N} &\equiv \mathbb{H} \cdot \mathbb{T} \cdot \mathbb{H} \\ &= \frac{1}{\sqrt{2}}\left[\begin{array}{c|c} 1 & 1 \\ \hline 1 & -1 \end{array}\right] \cdot \left[\begin{array}{c|c} C & S \\ \hline S & C \end{array}\right] \cdot \frac{1}{\sqrt{2}}\left[\begin{array}{c|c} 1 & 1 \\ \hline 1 & -1 \end{array}\right] \\ &= \frac{1}{2}\left[\begin{array}{c|c} C + S & S + C \\ \hline C - S & S - C \end{array}\right]\left[\begin{array}{c|c} 1 & 1 \\ \hline 1 & -1 \end{array}\right] \\ &= \left[\begin{array}{c|c} C + S & 0 \\ \hline 0 & C - S \end{array}\right] \\ &= \left[\begin{array}{c|c} e^{+TD} & 0 \\ \hline 0 & e^{-TD} \end{array}\right] \end{aligned}$$

or

$$\boxed{\mathbb{N} = \left[\begin{array}{c|c} \Delta_T^{-1} & 0 \\ \hline 0 & \Delta_T \end{array}\right]}$$

Thus we may write

$$\begin{array}{ccccc} \mathbb{R}_1 & = & \mathbb{N} & \cdot & \mathbb{R}_2 \\ \left[\begin{array}{c} u_1 \\ \hline v_1 \end{array}\right] & = & \left[\begin{array}{c|c} \Delta_T^{-1} & 0 \\ \hline 0 & \Delta_T \end{array}\right] & \cdot & \left[\begin{array}{c} u_2 \\ \hline v_2 \end{array}\right] \end{array}$$

where $\Delta_t \equiv e^{-TD}$ is the time delay operator. We may then obtain the two time domain relations:

$$\boxed{\begin{aligned} u_2(t) &= u_1(t - T) \\ v_1(t) &= v_2(t - T) \end{aligned}}$$

These characteristic relations were first obtained by Bernhard RIEMANN in 1860 for the nonlinear case of sound waves of finite amplitude.

Background Reading -- Method of Characteristics

(1) RIEMANN, Bernhard: Ueber die Fortpflanzung ebener Luftwellen von endlicher Schwingungsweite, (1860) Gesammelte Mathematische Werke, pp. 156-175, Second Edition (1953).

(2) MASSAU, Janius: Unsteady Flow in Open Channels, Annales de l'Association des Ingénieurs sortis des écoles spéciales de Gand T. 23 pp. 95-214 (1900).

(3) BERGERON, L.: Du Coup de Bélier en Hydraulique au Coup de Foudre en Electricité (1950)

Background Reading -- Scattering Matrices

(1) CARLIN, H. J.: The Scattering Matrix in Network Theory, IRE Transactions on Circuit Theory, Vol. CT-3, Number 2, pp. 88-97 (June, 1956).

(2) REDHEFFER, R. R.: Difference Equations and Functional Equations in Transmission-line Theory, Modern Mathematics for the Engineer, Second Series, pp. 282-337 (1961)

Input Impedance of a Lossless "TM" Driving a Linear "R" Load

System:

Structure:

$$E \frac{e_1}{f_1} \mathrm{TM} \frac{e_2}{f_2} R$$

Disturbance: $e_1(t) = E \sin \omega t$

Response: $f_1(t) = F \sin (\omega t - \phi)$

TRANSFER CHARACTERISTIC:

Setting the ratio $R_2/R_0 \equiv r$, the intrinsic transmission matrix yields:

$$\begin{bmatrix} \epsilon_1 \\ \hdashline \phi_1 \end{bmatrix} = \begin{bmatrix} C & \vdots & S \\ \hdashline S & \vdots & C \end{bmatrix} \cdot \begin{bmatrix} r\phi_2 \\ \hdashline \phi_2 \end{bmatrix}$$

Therefore the operational transfer characteristic becomes:

$$\mathbb{F}_{11} = \frac{\phi_1}{\epsilon_1} = \frac{S_r + C}{C_r + S} = \frac{r\mathcal{T} + 1}{r + \mathcal{T}} = \frac{r \tanh TD + 1}{r + \tanh TD}$$

FREQUENCY RESPONSE:

In terms of frequency response

$$\mathbb{F}_{11}(j\omega) = \frac{1 + j\, r \tan \omega T}{r + j \tan \omega T}$$ ----------(Periodic Functions)

Since the arguments are periodic, the transfer function will itself be periodic, Moreover, any transformation of the form $w = (a + bz)/(c + dz)$ with w and z complex numbers is a bilinear form which always transforms CIRCLES into CIRCLES (note that [St. LINES] $\subset$ [CIRCLES] and [INVERSE TF] $\subset$ [BILINEAR TF]).

In this case $z = j \tan \omega t$ is a straight line, and $\mathbb{F}_{11}(z)$ is therefore a circle. To determine the circle, it is only necessary to fix three points. Since all frequency responses of linear systems must give polar plots symmetric about the horizontal axis, we know that the center of the $\mathbb{F}_{11}$ circle must lie on the axis. The two additional points may be quickly found by noting:

(1) At: $\omega t = 0, \pm\pi, \pm 2\pi, \dots, \pm 2n\pi, \dots$

$$\tan \omega t \equiv 0$$

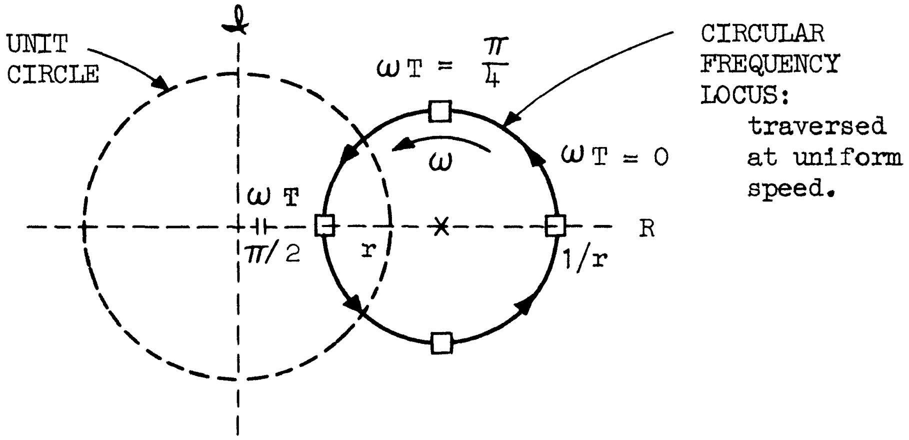
$$\boxed{F_{11} = 1/r}$$

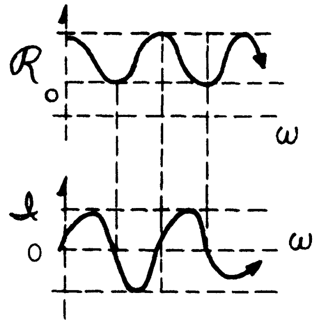
(2) At: $\omega t = \pm\frac{\pi}{2}, \pm\frac{3\pi}{2}, \dots, \pm\frac{2n+1}{2}\pi, \dots$

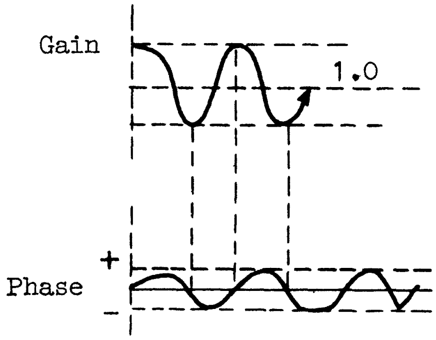
$$\tan \omega t \longrightarrow \infty$$

$$\boxed{F_{11} = r}$$

Thus the final locus has the appearance:

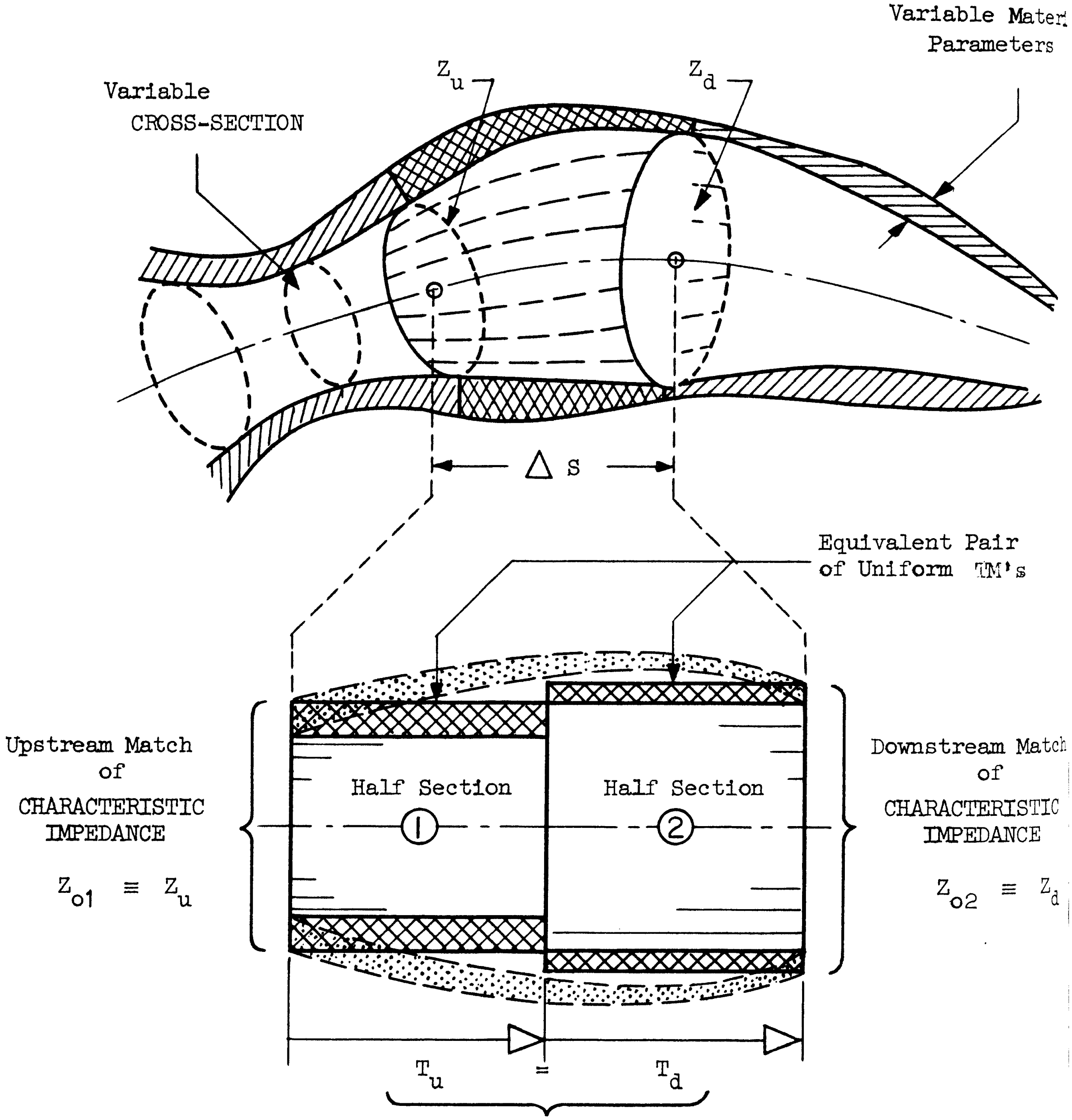






Nonuniform Transmitters

A general variable parameter wavelike transmitter can always be considered as a finite or infinite sequence of small elements, thus:



Each small section may be considered as two lengths of uniform transmitter having equal transmission times $T_u = T_d = T$. The properties and corresponding characteristic impedance of one half-section are taken as that at the beginning of the physical element while the parameters of the other half-section are those corresponding to the end of the physical element.

It is then apparent that as the lengths and corresponding transmission times shrink to zero this model will become exact, both for abrupt and for gradual non-uniformities.

The advantage of synchronous timing for all the elements can now be made evident. We may factor the transmission matrix for a double element into the intrinsic forms:

Ru, T
Upstream
Half-Section

Rd, T
Downstream
Half-Section

$$\left[\begin{array}{c|c} \sqrt{Ru} & 0 \\ \hline 0 & \sqrt{Gu} \end{array}\right] \left[\begin{array}{c|c} C & S \\ \hline S & C \end{array}\right] \left[\begin{array}{c|c} \sqrt{Gu} & 0 \\ \hline 0 & \sqrt{Ru} \end{array}\right] \left[\begin{array}{c|c} \sqrt{Rd} & 0 \\ \hline 0 & \sqrt{Rd} \end{array}\right] \left[\begin{array}{c|c} C & S \\ \hline S & C \end{array}\right] \left[\begin{array}{c|c} \sqrt{Gd} & 0 \\ \hline 0 & \sqrt{Rd} \end{array}\right]$$

$$\mathbb{T} \left[\begin{array}{c|c} \sqrt{Rd/Ru} & 0 \\ \hline 0 & \sqrt{Ru/Rd} \end{array}\right] \mathbb{T}$$

$$\mathbb{T} \left[\begin{array}{c|c} r & 0 \\ \hline 0 & 1/r \end{array}\right] \mathbb{T}$$

$$\mathbb{T} \cdot \quad \mathbb{E} \quad \cdot \mathbb{T}$$

TM — TF — TM

Thus it is that the matrix for the general case may be expressed (except for the terminal scaling) by the finite or infinite product:

$$\mathbb{M}_n = \prod_{k=1}^{n} \mathbb{T}^2 \, \mathbb{E}_k$$

In the special case where the transformer ratios, r_k, remains everywhere close to unity, the matrix $\mathbb{M}_n$ may be approximately refactored to the forms:

$$\mathbb{M}_n = (\mathbb{T}^{2n}) \left(\prod_{k=1}^{n} \mathbb{E}_k\right) \equiv \left(\prod_{k=1}^{n} \mathbb{E}_k\right) (\mathbb{T}^{2n})$$

Under these conditions, behavior is comparable to that of an equivalent uniform transmitter for which the local intrinsic variables

$$\epsilon(x) = \sqrt{G_o} \cdot e(x)$$
$$\phi(x) = \sqrt{R_o} \circ f(x)$$

satisfy the wave relations. This result is known as <u>Green's</u> <u>Law</u> after George GREEN, who in 1837 derived these relations for gravity waves in shallow water channels.

A practical example of the use of Green's Law for engineering estimates may be indicated in terms of the pressure surge following sudden shutoff in flow at the small end of a tapered pipe. The Green's Law estimate would be obtained from the formula:

$$P(s)/P_o \cong \sqrt{Z_o(s)/Z_{oo}} = \sqrt{D_o/D(s)}$$

where D is the pipe diameter. The following tabulation gives typical results

Station	Diameter	Calculated Pressure Rise	
		Exact Results	Green's Law
0	1.00	1.00	1.00
1	1.20	0.82	0.83
2	1.40	0.69	0.71
3	1.60	0.60	0.63
4	1.80	0.53	0.56
5	2.00	0.47	0.50

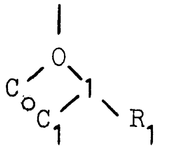
E. Modeling Diffusive Transmission

Wavelike Transmitters with Dispersion

Consider the transmission micro-element:

where $Y(d) \ddagger CD$ but is a general function of D. This situation is a much closer approximation to reality than the pure capacitive case.

For example, consider $Y(D) = C_o D + [C_1 D/(1 + T_1 D)]$ corresponding to the 1 port element:



$$T_1 = R_1 C_1$$

It is easy to demonstrate that this results in a dispersion or scattering action. Such dispersion is a vitally important factor in all communication and other transmission of information, since, for example it determines the effective "channel capacity" in the fundamental Shannon Formula:

$$C = W \ln [1 + (S/N)]$$

Let us consider a more general case in which $C_o = 0$ and

$$Y = \sum_k^n [C_k D/(1 + T_k D)]$$

Then

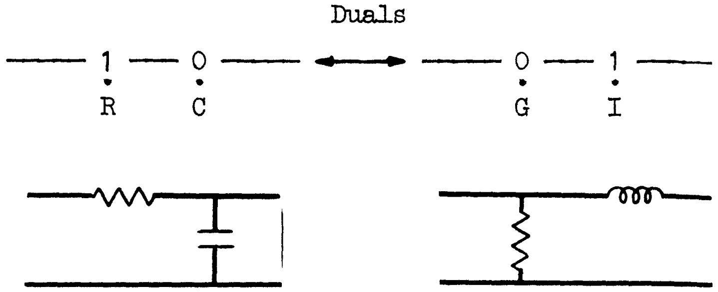
$$\text{for low frequencies:} \quad Y \rightarrow (\sum^n C_k) \cdot D = C_n D$$

$$\text{for high frequencies:} \quad Y \rightarrow \sum^n C_k/T_k = \sum G_k = G_n$$

This latter corresponds to a pure diffusive transmitter which was considered by William Thomson, Lord KELVIN, in connection with SUBMARINE CABLES.

Pure Diffusive Transmitters

Consider the micro elements:



This geometry first arose in J. B. J. FOURIER'S analysis of transient heat conduction: Theorie Analytique de Chaleur. The FOURIER mathematics was exhaustively applied by KELVIN to the RC cable with singular consequence

In our general TM parlance we now have:

$$\left.\begin{array}{l} Z_t = R_t \\ Y_t = C_t D \end{array}\right\} \therefore \left.\begin{array}{ll} \Gamma = \sqrt{Z_t \cdot Y_t} & = \sqrt{TD} \\ Z_o = \sqrt{Z_t / Y_t} & = R_t/\sqrt{TD} \end{array}\right\} T = R_t C_t$$

Note now that the characteristic impedance is an operator, no longer a scalar, while the time constant $T = R_t C_t$ is an R-C parameter. This is the microscopic counterpart of the distinction between restoring times and wave periods of macrosystems.

In order to interpret these universal properties let us consider the special case

------------ Relative Length ≡ 1 --------

$\frac{e_1}{f_1}$	Diffusive Transmitter	$\frac{e_x}{f_x}$	Diffusive Transmitter	$\frac{e_o}{f_o \equiv 0}$	Infinite Impedance

|←--- (1 — x)---→||←----(x) -----→|

Then:

$$\left[\begin{array}{c} e_x \\ \hdashline f_x \end{array}\right] = \left[\begin{array}{c:c} \mathbb{A}_x & \mathbb{B}_x \\ \hdashline \mathbb{C}_x & \mathbb{D}_x \end{array}\right] \cdot \left[\begin{array}{c} e_o \\ \hdashline 0 \end{array}\right]$$

$$\left[\begin{array}{c} e_1 \\ \hdashline f_1 \end{array}\right] = \left[\begin{array}{c:c} \mathbb{A}_1 & \mathbb{B}_1 \\ \hdashline \mathbb{C}_1 & \mathbb{D}_1 \end{array}\right] \cdot \left[\begin{array}{c} e_o \\ \hdashline 0 \end{array}\right]$$

This gives

$$\boxed{\frac{e_x}{e_1} = \frac{\mathbb{A}_x}{\mathbb{A}_1} = \frac{\cosh x\sqrt{TD}}{\cosh\sqrt{TD}}}$$

Then for example at the "insulated" end (x = 0)

$$e_x(t) = \left[\frac{1}{\cosh\sqrt{TD}}\right] \cdot e_1(t) = \mathbb{F}(D) \cdot e_1(t)$$

Here as in all such diffusive systems, the transfer operator $\mathbb{F}(D)$ is a monotone process which bears to diffusion the corresponding relation that vibratory processes bear to wavelike transmission. This we shall treat next.

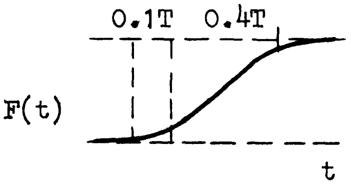
However, first we may interpret our result by factoring the "cosh" function into its characteristic roots. This yields a representation as an infinite set of tapered first order lags.

$$\mathbb{F}(D) = \frac{1}{\cosh\sqrt{TD}} = \prod_{k=1}^{\infty}\left[\frac{1}{1 + a_k TD}\right] \quad \text{where } \partial a_k = \frac{4/\pi^2}{(2k-1)^2}$$

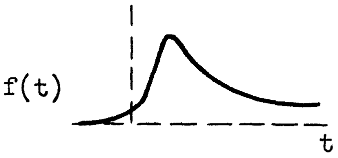
$$= \frac{1}{1 + 0.40\ TD} \cdot \frac{1}{1 + 0.04\ TD} \cdot \frac{1}{1 + 0.02\ TD} \cdots$$

resulting in the responses:

STEP RESPONSE:



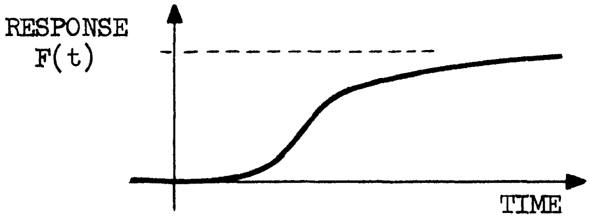
IMPULSE RESPONSE:



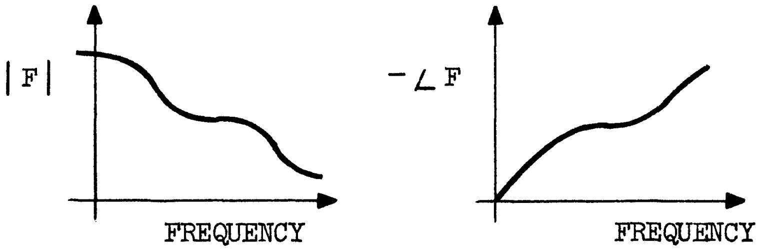
F. The Dynamics of Monotone Processes

Introduction

A large number of fluid, thermal, chemical and other industrial and organic processes are characterized by a step response which is monotonic non-decreasing in time as indicated.



The corresponding frequency response, at least for most continuous processes, would have a non-increasing amplitude and non-decreasing phase lag with increasing frequency as follows:



All linear systems giving rise to such response can be called monotone processes.

Monotone response is manifested through:

(a) Time delay or dead time.

(b) Dispersion or rise time.

In physical processes, time delay is usually associated with propagation or transport phenomena as measured by the ratio of travel distance to propagation or transport velocity. The dispersion in any process can ultimately be attributed to the law of increasing entropy, whereby the distributed resistances in any system cause an attenuation increasing with frequency. Such scattering action is reduced by isolation and relaying methods, but is always present to some degree.

Oscillatory processes, characterized by the presence of complementary energy storage elements, will have monotonic response whenever the energy dissipated per cycle becomes sufficiently large compared to the energy stored in each mode.

Operational Transforms

One finds recurrent need to define quantities which are definite integrals with respect to a given monotone distribution function F(t). Thus we may write the expectation of a function $\mathbb{E}(g)$, as

Expectation: $$\mathbb{E}(g) \equiv \int_{-\infty}^{+\infty} g \cdot dF(t)$$

Thus one finds that the moments about the origin of a monotone may be expressed compactly in the Stieltjes form by the relations:

k-th moment: $$a_k = \mathbb{E}(t^k) = \int_{-\infty}^{+\infty} t^k \cdot dF(t)$$

The infinite set of such moments, of course, measures the distribution properties of a monotone function; in particular, the zero moment:

$$a_o = \int_{-\infty}^{+\infty} dF(t)$$

measures the total area of the distribution. One conventionally normalizes the distribution, if possible, such that a_o is identically unity to give $F(-\infty) \equiv 0$ and $F(\infty) \equiv 1$. We shall henceforth assume this to be the case.

The first moment:

$$a_1 = \int_{-\infty}^{+\infty} t \cdot dF(t)$$

yields the mean effective position or centroid of the distribution if $a_o = I$.

In the same way the second moment:

$$a_2 = \int_{-\infty}^{+\infty} t^2 \cdot dF(t)$$

Measures the mean square position of the distribution, and so forth.

In addition, the (bilateral) Laplace transform $\mathbb{F}(s)$ of a mono-Laplace transform:

$$\mathbb{F}(s) \equiv \mathbb{E}(e^{-st}) = \int_{-\infty}^{+\infty} e^{-st} \cdot dF(t)$$

while the corresponding Fourier transform, $\mathbb{F}(\omega)$, becomes Fourier transform:

$$\mathbb{F}(\omega) \equiv \mathbb{E}(e^{-j\omega t}) = \int_{-\infty}^{+\infty} e^{-j\omega t} \cdot dF(t)$$

In general, for monotone functions, the above integral transform functions all exist in the mathematical sense, and uniquely characterize the given function. It is readily understood that all physically realizable dynamic monotone responses must satisfy the condition.

$$F(t) \equiv o \quad \text{for} \quad t < o$$

since otherwise an effect would occur in the absence of cause, which is a situation not normally encountered.

Moreover, certain relationships exist between the above transforms. For example, if in the equation above, the exponential e^{-st} is expanded in an infinite power series,

$$e^{-st} = 1 - st + \frac{1}{2} s^2t^2 - \cdots = \sum_{k=o}^{\infty} \frac{(-s)^k}{k!} t^k$$

and the result is integrated term-by-term, there results the series expansion for the general monotone operator, namely:

$$\mathbb{F}(s) = \sum_{k=o}^{\infty} \frac{a_k(-s)^k}{k!} t^k$$

Thus, the Laplace transform of any monotone distribution function is very simply expressed as an alternating power series whose coefficients are directly related to the moments about the origin of its original (or time) distribution.

In the same way, we may also write the Fourier transform of the distribution in terms of a power series which may be found simply by placing $s = j\omega$ in the expansion above to obtain:

$$\mathbb{F}(\omega) = \sum_{k=o}^{\infty} \frac{a_k(-j\omega)^k}{k!}$$

If we separate this series into its real and imaginary parts, there results

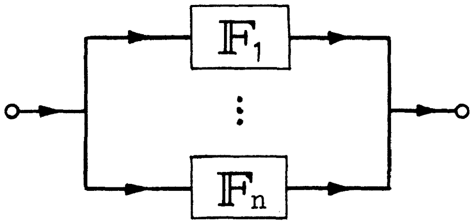
$$\operatorname{Re} \mathbb{F}(\omega) = a_o - \frac{a_2}{21}\omega^2 + \frac{a_4}{4!}\omega^4 - \cdots$$

$$-\operatorname{Im} \mathbb{F}(\omega) = a_1\omega - \frac{a_3}{3!}\omega^2 + \frac{a_5}{5!}\omega^5 - \cdots$$

which means that the real and imaginary components of the frequency response of any monotone process are given by simple alternating power series expansions of even and odd powers, whose coefficients are directly proportional to the corresponding moments of the step response. This fact offers one possibility for determining the moments, and therefore the transient response characteristics directly from observed frequency response data.

Parallel Combination of Monotones

The operational sum of a sequence of monotone responses is itself a monotone response. Such a situation arises whenever two or more monotone processes are placed in parallel as indicated.



This situation is expressed by the operational equation:

$$\mathbb{F}_p(s) = \mathbb{F}_1 + \mathbb{F}_2 + \cdots + \mathbb{F}_n = \sum_{m=1}^{n} \mathbb{F}_m(s)$$

Therefore the moments, a_{kp}, of the resultant distribution are given by:

$$a_{kp} = \sum_{m=1}^{n} a_{km}$$

Cascade Combination of Monotones

The operational product of a sequence of monotone responses is itself a monotone response. Such a situation occurs whenever two or more monotone processes are put in tandem or cascade as indicated.

This circumstance is represented by the operational equation:

$$\mathbb{F}_c(s) = \mathbb{F}_1 \cdot \mathbb{F}_2 \cdot \ldots \cdot \mathbb{F}_n = \prod_{m=I}^{n} \mathbb{F}_m(s)$$

The logarithms of the $\mathbb{F}_m$, will, however, add, in the form:

$$\log \mathbb{F}_c = \sum_{m=1}^{n} \log \mathbb{F}_m$$

The transmission operator, log $\mathbb{F}$, corresponding to any monotone $\mathbb{F}$, can also be expanded in a power series of the form

$$\log \mathbb{F} = \sum_{k=o}^{\infty} \frac{c_k}{k!} (-s)^k$$

where the coefficients c_k are called (by statisticians) the cumulants or semi-invariants of the distribution F(t).

Thus, for any cascade of n monotones, the cumulants are additive in the form

$$c_{kc} = \sum_{m=1}^{n} c_{km}$$

The moments a_k and corresponding cumulants c_k are related by the identity:

$$a_k = c_k + \sum_{m=I}^{k-1} \binom{k-1}{m-1} c_m a_{k-m}$$

Since c_k has the dimensions of a_k and therefore t^k, it is convenient and significant to define the following set of constants:

Attenuation	δ	$\equiv$	$\ln c_o$
Mean Delay	T_m	$\equiv$	c_1
Dispersion Time	T_s	$\equiv$	$(c_2)^{1/2}$
Skew Time	T_a	$\equiv$	$(c_3)^{1/3}$
Excess Time	T_e	$\equiv$	$(c_4)^{1/4}$

In terms of these new constants any monotone process may be characterized by the transform:

$$F(s) = e^{\delta - T_m s + \frac{1}{2} T_s^2 s^2 - \frac{1}{6} T_a^3 s^3 + \frac{1}{24} T_e^4 s^4 - \cdots}$$

In terms of the frequency response, this expansion demonstrates that the amplitude depends only on the even powers of ω and the phase only on the odd powers since

Gain

$$\log |\mathbb{F}| = \delta - \frac{1}{2} T_s^2 \omega^2 + \frac{1}{24} T_e^4 \omega^4 - \cdots$$

Phase

$$\angle \mathbb{F} = - T_m \omega + \frac{1}{6} T_a \omega^3 - \cdots$$

Thus, the description of a monotone process is unique only if an infinite set of parameters is specified. However, any monotone may be approximated with increasing accuracy by matching an increasing number of the cumulants of the actual process by their counterparts in a model.

One may also profitably introduce into dynamic context, the conventional statistical dimensionless coefficients:

Coefficient of Variance $\mu \equiv T_s/T_m$

Coefficient of Skew $\alpha \equiv T_a^3/T_s^3$

Coefficient of Excess $\beta \equiv T_e^4/T_s^4$

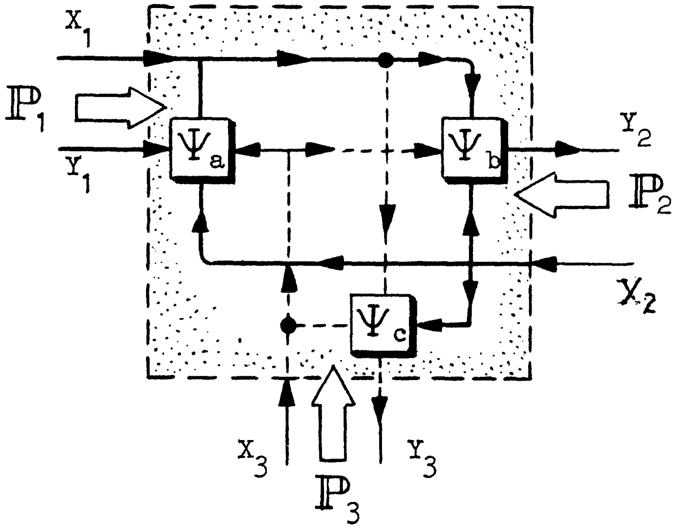
XVII. Energy Modulation and Amplification

A. General Three-Port Elements

The causal relationships between the three inputs and the three outputs of a general three-port element are of the form:

$$Y_1(t) = \Psi_a \, [X_1(t) \quad X_2(t) \quad X_3(t)]$$
$$Y_2(t) = \Psi_b \, [X_1(t) \quad X_2(t) \quad X_3(t)]$$
$$Y_3(t) = \Psi_c \, [X_1(t) \quad X_2(t) \quad X_3(t)]$$

which would be diagrammed as follows:



For static behavior the Ψ's reduce to Φ's, while for linearized dynamic response the Ψ's partition in the form:

$$Y_1 = \mathbb{F}_{11} \cdot X_1 + \mathbb{F}_{12} \cdot X_2 + \mathbb{F}_{13} \cdot X_3$$
$$Y_2 = \mathbb{F}_{21} \cdot X_1 + \mathbb{F}_{22} \cdot X_2 + \mathbb{F}_{23} \cdot X_3$$
$$Y_3 = \mathbb{F}_{31} \cdot X_1 + \mathbb{F}_{32} \cdot X_2 + \mathbb{F}_{33} \cdot X_3$$

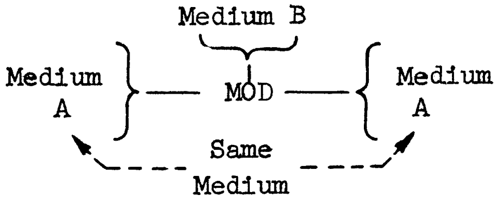
which is the particularization for 3-ports of the linear relation:

$$\mathbb{Y} = \boldsymbol{\Lambda} \cdot \mathbb{X}$$

B. Generalized Power Modulators as Three-Port Elements

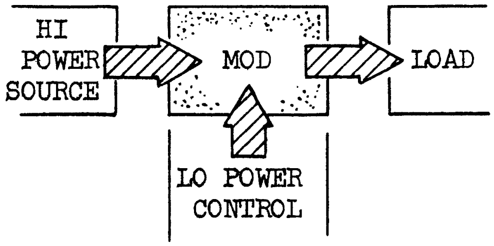
Perhaps the most significant single development of the twentieth century has been the sharpening of the concepts and practice concerned with the <u>modulation</u> and <u>amplification</u> of power and signals. The primitive elements required for all such transformations involve significant energy transfer at a minimum of three ports. Thus the three-port element serves as the prototype <u>generalized</u> <u>modulator</u>.

We are here concerned with modulators in all media (i.e., electrical fluid, mechanical, thermal, and so forth), but it is important to emphasize that for a three-port to be considered as a true power modulator it will generally have the bond pattern indicated:



While it is often possible to have Medium B ≡ Medium A, one would not usually consider three-ports involving three different media as true modulators. Some typical species and realizations are indicated in the morphological matrix of Table I.

Thus the normal power flow for a modulator would appear as follows:



Moreover, it is very convenient to have a canonical ordering of the ports as follows:

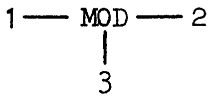


Table I

SPECIES OF MODULATORS

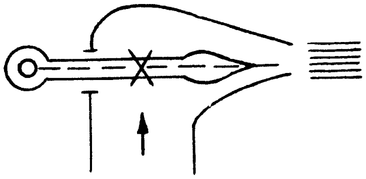
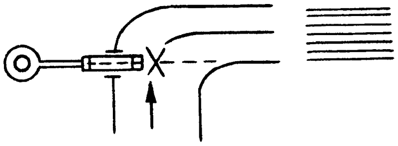
A Morphological Matrix of Certain Realizations

Principal Medium A	Modulating Medium B			
	Fluid	Electrical	Mechanical	Thermal
Fluid	Pneumatic or Hydraulic Operated Valves	Electrically Operated Valves MHD Devices	VALVES OF ALL TYPES	
Electrical		VACUUM TUBES TRANSISTORS Saturable Reactors	SWITCHES RELAYS	Bimetal Switches SUPER-CONDUCTIVE RELAYS
Mechanical	Modulated FLUID COUPLING	Magnetic Clutch	CLUTCHES DIFFERENTIALS Variable Speed Drives	
Thermal	BISURFACE MODULATED CONVECTOR			

for which under normal steady conditions the flow of power would be $1 \longrightarrow 2$ with the inequalities

$$\mathbb{P}_1 \geq \mathbb{P}_2 >> \mathbb{P}_3$$

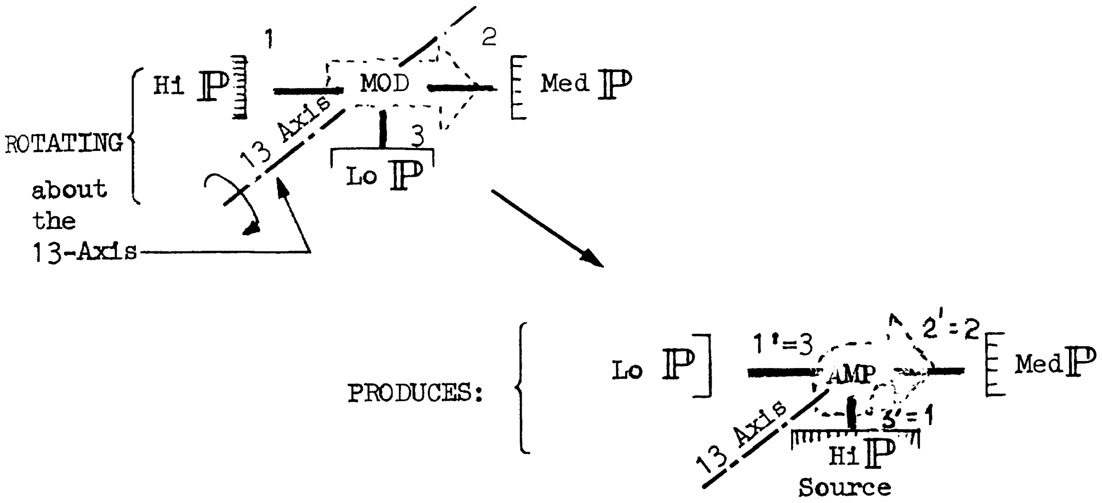
Two very broad classes of modulators exist according to whether the above weak inequality becomes a weak equality or strong inequality. These may be considered as follows:

WEAK EQUALITY	STRONG INEQUALITY
$\mathbb{P}_1 \approx \mathbb{P}_2$	$\mathbb{P}_1 > \mathbb{P}_2$
STRUCTURAL or PARAMETRIC MODULATORS - - - - - - - Modulate by TRANSFORMER ACTION - - - - - - - HIGH EFFICIENCY	DISSIPATIVE or THROTTLING MODULATORS - - - - - - - Modulate by CONVERTING $\mathbb{P} \rightarrow P_d$ - - - - - - LOW EFFICIENCY
Prototype Example: NEEDLE VALVE 	Prototype Example: GATE VALVE 

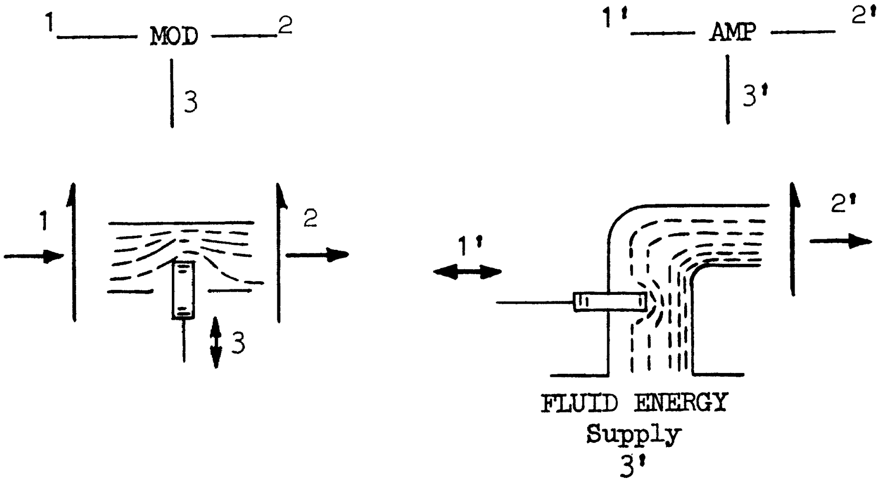
Henceforth we shall find it convenient for visualization purposes to take the fluid valve as a prototype modulator and the gate valve as the most representative type.

C. Generalized Amplifiers

Any modulator can be used as an <u>amplifier</u> by the elementary transposition of configuration indicated:

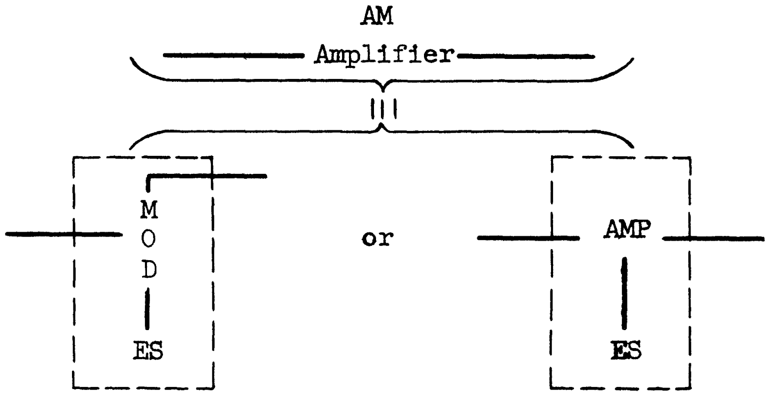


This is readily seen in terms of the gate valve:



The modulator (MOD) and the amplifier (AMP) can thus be considered as perturbations of the same primitive device, merely resulting from interchange of supply and control ports.

Frequently an amplifier is viewed as an active (i.e., nonreciprocal) two-port in the fashion:



Where "ES -" represents an energy source.

Thus a cascade or chain of amplifiers in any medium could be represented by the bond diagram:

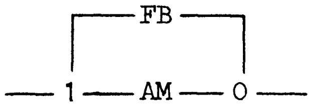
... ——AM——TP——AM——TP——AM—— ...

where ——TP—— represents any two-port coupling system.

Similarly the use of EFFORT FEEDBACK or FLOW FEEDBACK around any amplifier can be indicated as follows, using ——FB—— for the two-port feedback elements:

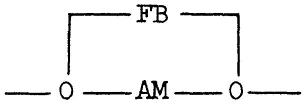
EFFORT FEEDBACK

(Voltage, Force, Pressure, etc.)



FLOW FEEDBACK

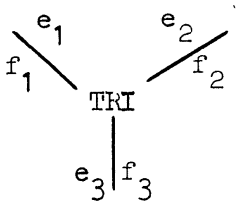
(Current, Motion, Flow, etc.)



These cases will be discussed further below.

D. The Trinode as a Three-Port Element

The following relations are true for any three-port:



$$f_1 = \Psi_1(e_1, e_2, e_3) \qquad \ldots \quad \text{I}$$

$$f_2 = \Psi_2(e_1, e_2, e_3) \qquad \ldots \quad \text{II}$$

$$f_3 = \Psi_3(e_1, e_2, e_3) \qquad \ldots \quad \text{III}$$

But frequently in addition there exist additional constraints upon the effort and flow variables.

For example, if all ports connect the same medium, it is reasonable to expect that the continuity equation will hold for the three flows, namely, for positive inward flows:

$$f_1 + f_2 + f_3 \equiv 0 \qquad \ldots \quad \text{IV}$$

Moreover, any three-port which depends only upon relative or differential efforts will obey the condition for any constant effort, E:

$$\Psi((e_1 + E), (e_2 + E), (e_3 + E)) = \Psi(e_1, e_2, e_3) \qquad \ldots \quad \text{V}$$

for all three flow functionals.

Such a three-port we shall find useful to recognize and denote as a <u>trinode</u>. Particularly noteworthy are the particular instances of vacuum triodes and transistors.

If the trinode be linear, we could write a linear admittance matrix in the form:

$$\mathbf{f} = \mathbb{Y} \cdot \mathbf{e}$$

$$\left[\begin{array}{c} f_1 \\ \hdashline f_2 \\ \hdashline f_3 \end{array}\right] = \left[\begin{array}{c:c:c} Y_{11} & Y_{12} & Y_{13} \\ \hdashline Y_{21} & Y_{22} & Y_{23} \\ \hdashline Y_{31} & Y_{32} & Y_{33} \end{array}\right] \cdot \left[\begin{array}{c} e_1 \\ \hdashline e_2 \\ \hdashline e_3 \end{array}\right]$$

But the continuity condition (IV) further requires that the sum of each column vanish:

$$\sum_i Y_{ij} = 0 \qquad \ldots (IV')$$

while the datum relativity condition (V) requires that each row sum vanish:

$$\sum_j Y_{ij} = 0 \qquad \ldots (V')$$

These results were first pointed out by SHEKEL and mean that if we consider each of the constraints upon the general three-port matrix, we would find the following:

GENERAL Three-port	9 Elements
Less Continuity	-3
Less Relativity	-2
ACTIVE Trinode	4 Elements
Less Reciprocity	-1
PASSIVE Trinode	3 Elements
Less Symmetry	-2
UNIFORM Trinode (flow Jct.)	1 Element

Thus it is clear then that the principal difference between, on the one hand, a linear active trinode, with internal energy sources, and, on the other hand, passive elements, lies in the failure of the former element to satisfy the reciprocity conditions.

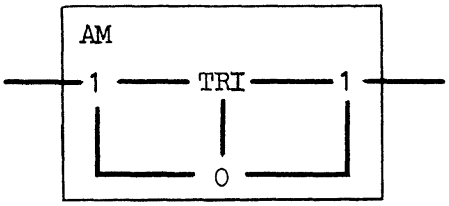
Trinode Amplifiers

The practical significance of the active or nonreciprocal trinode modulating element lies in its value as an amplifier or relay. This will generally be of the two-port form:

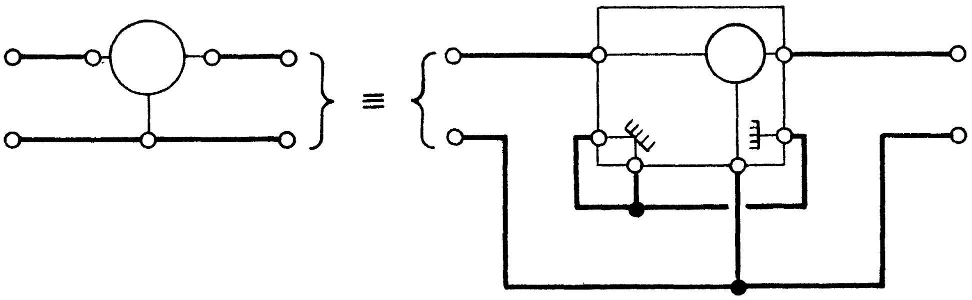
—— AM ——

or: ——>>——

which may be realized from any active trinode by inserting the trinode into the junction structure:



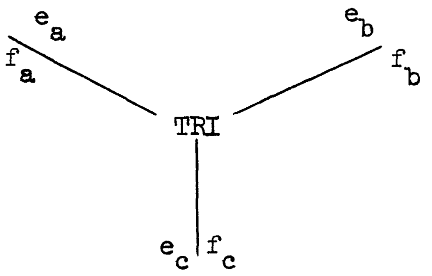
This is, of course, merely a generalization of the ordinary electric circuit configuration:



We may then derive the various configurational permutations in an entirely systematic way as shown below.

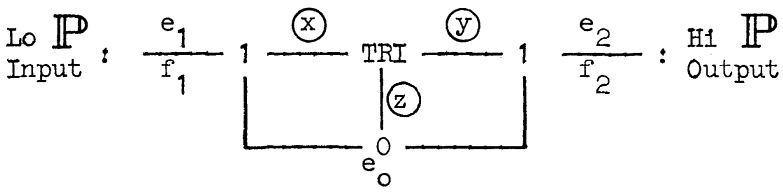
Trinode Amplifier Configurations

The admittance matrix for a generalized linear trinode would have the form:



$$\mathbb{Y} = \begin{array}{ccc|c} a & b & c & \\ \hline Y_{aa} & Y_{ab} & Y_{ac} & a \\ Y_{ba} & Y_{bb} & Y_{bc} & b \\ Y_{ca} & Y_{cb} & Y_{cc} & c \end{array}$$

If we consider Port-a as the low power input, there exist three amplifier configurations all of the general form:



The effort e_o is common to both sides of the amplifier (——>>——). The permutations are as indicated:

	Common-a	Common-b	Common-c
Port x :	c	a	a
Port y :	b	c	b
Port z :	a	b	c

All three of the two-port amplifier configurations may now be determined from the general results:

Two-Port ADMITTANCE Matrix : $\mathbb{Y}_{xy} \equiv \left[\begin{array}{c|c} Y_{xx} & Y_{xy} \\ \hline Y_{yx} & Y_{yy} \end{array}\right]$

Determinant : $\Delta_{xy} \equiv Y_{xx}Y_{yy} - Y_{xy}Y_{yx}$

Two-Port TRANSMISSION: Matrix $\mathbb{M}_{xy} \equiv \frac{-1}{Y_{yx}}\left[\begin{array}{c|c} Y_{yy} & 1 \\ \hline \Delta_{xy} & Y_{xx} \end{array}\right]$

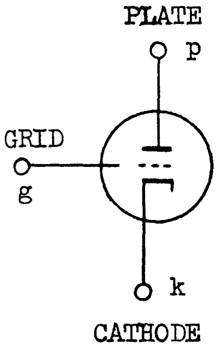
The corresponding matrices and matrices for each particular case are as follows:

CONFIGURATION:	Common-a	Common-b	Common-c
STRUCTURE:	c ▼ b a	a ▼ c b	a ▼ b c
$\mathbb{Y}$:	$\left[\begin{array}{c\|c} Y_{cc} & Y_{cb} \\ \hline Y_{bc} & Y_{bb} \end{array}\right]$	$\left[\begin{array}{c\|c} Y_{aa} & Y_{ac} \\ \hline Y_{ca} & Y_{cc} \end{array}\right]$	$\left[\begin{array}{c\|c} Y_{aa} & Y_{ab} \\ \hline Y_{ba} & Y_{bb} \end{array}\right]$
$\mathbb{M}$:	$\frac{-1}{Y_{bc}}\left[\begin{array}{c\|c} Y_{bb} & 1 \\ \hline \Delta_{cb} & Y_{cc} \end{array}\right]$	$\frac{-1}{Y_{ca}}\left[\begin{array}{c\|c} Y_{cc} & 1 \\ \hline \Delta_{ac} & Y_{aa} \end{array}\right]$	$\frac{-1}{Y_{ba}}\left[\begin{array}{c\|c} Y_{bb} & 1 \\ \hline \Delta_{ab} & Y_{aa} \end{array}\right]$

The Trinode Admittance of a Vacuum Triode

(Neglecting Interelectrode Capacitance)

This commonly used trinode element has the following standard symbol and linear admittance:



$$\mathbb{Y} = \begin{array}{ccc|c} g & p & k & \\ 0 & 0 & 0 & g \\ g_m & g_p & -(g_m + g_p) & p \\ -g_m & -g_p & g_m + g_p & k \end{array}$$

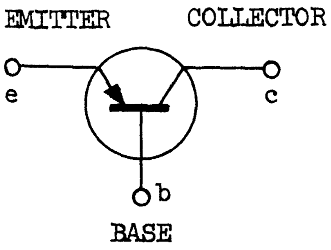
where $g_p = 1/r_p$ = Plate Conductance

$g_m = \mu g_p$ = Grid · Plate Transconductance

The Trinode Admittance of a Junction Transistor

(Neglecting all but first order effects)

This commonly used element has the following symbol and approximate admittance:



$$\mathbb{Y} = \begin{array}{ccc|c} e & c & b & e \\ g_i & 0 & -g_i & e \\ -\alpha g_i & g_o & +\alpha g_i - g_o & c \\ (\alpha - 1)g_i & -g_o & (1 - \alpha)g_i + g_o & b \end{array}$$

where $\alpha = \alpha$ (D) = Current Amplification Factor

g_i = Input Conductance $\sim g_b$

g_o = Output Conductance $\sim g_c \sim 0$

The Activity of Two-Ports

A useful measure of the "activeness" or activity of a non-reciprocal two-port derives from the determinant of the governing transmission matrix. From the previous results we have:

$$\Delta = Y_{xy}/Y_{yx}$$

If the element or system is reciprocal

$$Y_{xy} \equiv Y_{yx}$$

and

$$\Delta \equiv 1$$

Since, physically, Y_{xy} measures $\partial f_x / \partial e_y$ and $X_{yx} = \partial f_y / \partial e_x$, an amplifier with near-infinite power gain would have

$$\Delta \longrightarrow 0$$

This is true for the idealized triodes and transistors usually considered.

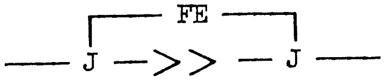
Then if we measure the activity of a two-port system as the unity complement of the absolute value of the determinant we have, using the Hebrew letter aleph:

$$\boxed{\text{Activity } \aleph \equiv 1 - |\Delta|}$$

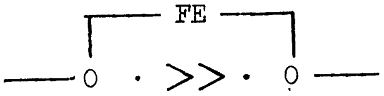
E. Cascading and Feedback of Amplifiers

Dynamics of Amplifier and Control Elements

If we connect across an amplifier (—>>—) another two-port element, using either 0 or 1 junctions (indicated by J), we arrive at the commonly encountered feedback scheme:

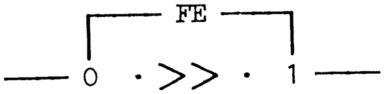


The two-port " —— FE —— " indicates the subsystem of particular feedback elements employed. The four particularizations of the energy junction situation may be denoted as follows (with typical electrical usage also indicated).



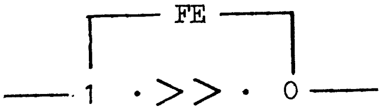
Flow-Flow Feedback
(SHUNT-SHUNT)

$$\mathbb{Y} = \mathbb{Y}_a + \mathbb{Y}_f$$



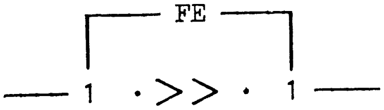
Flow-Effort Feedback
(SHUNT-SERIES)

$$\mathbb{G} = \mathbb{G}_a + \mathbb{G}_f$$



Effort-Flow Feedback
(SERIES-SHUNT)

$$\mathbb{H} = \mathbb{H}_a + \mathbb{H}_f$$



Effort-Effort Feedback
(SERIES-SERIES)

$$\mathbb{Z} = \mathbb{Z}_a + \mathbb{Z}_f$$

The Dynamics of Amplifier Chains

Consider a chain of amplifiers with the reticulation:

[—— AC ——] = [—— AM —— CN —— $]^n$ = [—— AS —— $]^n$

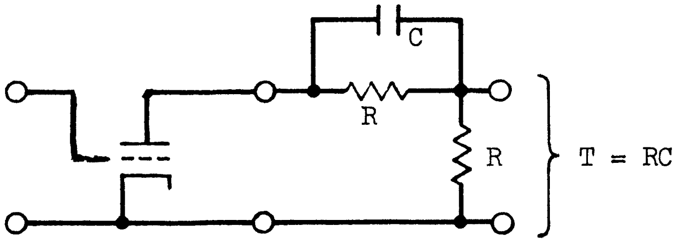
where [—— CN ——] is the interstage coupling network. If the elements are all linear, each stage may be represented in the form:

$$[\,-\,\mathrm{AS}\,-\,] \quad = \quad [\,-\,\mathrm{AM}\,-\,] \quad \cdot \quad [\,-\,\mathrm{CN}\,-\,]$$

$$\left[\begin{array}{c:c} \mathbb{A}_s & \mathbb{B}_s \\ \hdashline \mathbb{C}_s & \mathbb{D}_s \end{array}\right] = \left[\begin{array}{c:c} \mathbb{A}_a & \mathbb{B}_a \\ \hdashline \mathbb{C}_a & \mathbb{D}_a \end{array}\right] \cdot \left[\begin{array}{c:c} \mathbb{A}_c & \mathbb{B}_c \\ \hdashline \mathbb{C}_c & \mathbb{D}_c \end{array}\right]$$

We might take the vacuum tube common-cathode reactively coupled amplifier as an example. Here the single stage, — AS — , is given by:

$$\text{——— AM ———} \quad \cdot \quad \text{——— CN ———}$$



$$\left[\begin{array}{c:c} -1/\mu & r_p/\mu \\ \hdashline 0 & 0 \end{array}\right] \cdot \left[\begin{array}{c:c} \dfrac{2 + TD}{1 + TD} & \dfrac{R}{1 + TD} \\ \hdashline 1/R & 1 \end{array}\right]$$

$$\left[\begin{array}{c:c} \dfrac{(2 + TD)}{\mu(1 + TD)} + \dfrac{r_p}{\mu R} & \dfrac{-\,R}{\mu(1 + TD)} + \dfrac{r_p}{\mu} \\ \hdashline 0 & 0 \end{array}\right]$$

The voltage amplification ratio for a chain of "n" such amplifiers is therefore:

$$\frac{E_n(t)}{E_o(t)} = (1/\mathbb{A})^n = (-K)^n \left(\frac{1 + TD}{1 + kTD}\right)^n$$

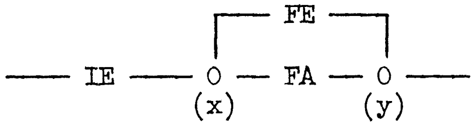
where $K = \mu/(2 + f)$

$k = (1 + f)/(2 + f)$

$f = r_p/\mu R$

Computing or Operational Amplifiers

A very useful and significant system configuration results whenever an ultrahigh gain, low-pass, inverting amplifier is employed with additional input and feedback elements, according to the scheme:



where: FA : Feedback Amplifier
FE : Additional Feedback Elements
IE : Input Elements

The admittance matrices of (— FE —) and (— FA —) will add, since they are coupled "flow-to-flow."

$$\mathbb{Y}_{FA} \cong \left[\begin{array}{c|c} 0 & 0 \\ \hline -G & 0 \end{array}\right]; \qquad \mathbb{Y}_{FE} = \left[\begin{array}{c|c} Y_{xx} & Y_{xy} \\ \hline Y_{yx} & Y_{yy} \end{array}\right]$$

$$\therefore \quad \mathbb{Y}_{amp} = \mathbb{Y}_{FA} + \mathbb{Y}_{FE} = \left[\begin{array}{c|c} Y_{xx} & Y_{xy} \\ \hline Y_{yx} - G & Y_{yy} \end{array}\right]$$

From this follows the transmission matrix, assuming $G \to \infty$:

$$\mathbb{M}_{amp} \cong \left[\begin{array}{c|c} 0 & 0 \\ \hline -Y_{xy} & 0 \end{array}\right]$$

Then, given any two-port (— IE —), with the matrix:

$$\mathbb{M}_{i} = \left[\begin{array}{c|c} \mathbb{A}_i & \mathbb{B}_i \\ \hline \mathbb{C}_i & \mathbb{D}_i \end{array}\right]$$

the overall system matrix can be determined as

$$\mathbb{M} = \mathbb{M}_i \quad \mathbb{M}_{amp}$$

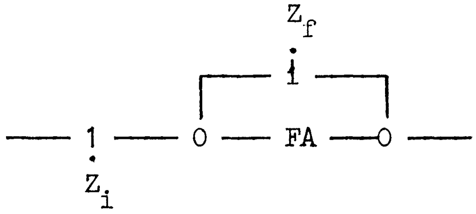
$$= \left[\begin{array}{c|c} \mathbb{A}_i & \mathbb{B}_i \\ \hline \mathbb{C}_i & \mathbb{D}_i \end{array}\right] \left[\begin{array}{c|c} 0 & 0 \\ \hline -Y_{xy} & 0 \end{array}\right]$$

$$= \left[\begin{array}{c|c} --\mathbb{B}_i Y_{xy} & 0 \\ \hline -\mathbb{C}_i Y_{xy} & 0 \end{array}\right]$$

The overall voltage transfer ratio is then, simply

$$E_2/E_1 = 1/\mathbb{B}_i Y_{xy} = - Z_{xy}/\mathbb{B}_i$$

If we take the very simple configuration:



then the resulting transfer ratio is

$$\boxed{E_2/E_1 = - Z_f/Z_i}$$

This structure, developed about twenty years ago in connection with gun directors and fire control apparatus, is the primitive element underlying the contemporary electronic analog computing machine or "electronic differential analyzer". The amplifiers (— FA —) developed for such use are referred to as d.c. computing amplifiers or operational amplifiers. Standardized input and feedback circuits permit the realization of scaling, summing, integrating, and other operations, as indicated in the voluminous literature on modern electronic analog computers.

Background Reading -- Matrix Methods for Active Systems

(1) ABBOTT, W. R.: Analysis of Four-Terminal Networks Containing Vacuum Tubes, Misc. Paper 46-204, AIEE (September, 1946).

(2) PETERSON, L. C.: Equivalent Circuits of Linear Active Four-Terminal Networks, The Bell Systems Technical Journal, Vol. XXVII, No. 4 pp. 593-622 (October, 1948).

(3) BROWN, J. S. and BENNETT, F. D.: The Application of Matrices to Vacuum-Tube Circuits, Proc. IRE, Vol. 36, pp. 844-852 (1948)

(4) EPSTEIN, H.: Solution of Transients in Active Four-Terminal Networks, J. Franklin Institute, Vol. 251, pp. 607-616 (1951)

(5) HSU, H.: On Transformations of Linear Active Networks with Applications at Ultra-High Frequencies, Proc. IRE, Vol. 41, pp. 59-67 (1953)

The five papers above were principally responsible for the introduction of 2-port matrix techniques to the design of vacuum-tube and transistor circuits.

(6) MIDDLEBROOK, R. D.: An Introduction to Junction Transistor Theory (1957)

(7) SHEA, R. F., Editor: Transistor Circuit Engineering (1957)

(8) -------------------: Principles of Transistor Circuits (1953)

These three books amply testify to the value of linear and nonlinear 2-port concepts in the design and applications of solid state amplifiers.

(9) WEBER, Ernst: Linear Transient Analysis, Vol. II (1956)

An excellent summary of much of the material in the above sources.