

AFCRL - 69-0273

692120

On a Physical Foundation for Electromagnetic
System Theory

Donald R. Rhodes
North Carolina State University
Raleigh, North Carolina 27607

Contract No. F 19628-67-C-0255
Project No. 4600
Task No. 460010
Work Unit No. 46001001

Scientific Report No. 1
20 June 1969

Distribution of this document is unlimited. It may be released to the Clearinghouse, Department of Commerce, for sale to the general public.

Contract Monitor: Allan C. Schell
Microwave Physics Laboratory

Prepared
for

Air Force Cambridge Research Laboratories
Office of Aerospace Research
United States Air Force
Bedford, Massachusetts 01730

Reproduced by the
CLEARINGHOUSE
for Federal Scientific & Technical
Information Springfield Va 22151

AFCRL - 69-0273

On a Physical Foundation for Electromagnetic
System Theory

Donald R. Rhodes
North Carolina State University
Raleigh, North Carolina 27607

Contract No. F 19628-67-C-0255
Project No. 4600
Task No. 460010
Work Unit No. 46001001

Scientific Report No. 1
20 June 1969

Distribution of this document is unlimited. It may be released to the Clearinghouse, Department of Commerce, for sale to the general public.

Contract Nomitor: Allan C. Schell
Microwave Physics Laboratory

Prepared
for

Air Force Cambridge Research Laboratories
Office of Aerospace Research
United States Air Force
Bedford, Massachusetts 01730

ABSTRACT: The traditional basis for the theory of electromagnetic systems (e.g., circuits, microwave networks, antennas) is a collection of definitions describing their temporal and spectral properties. It is shown here that these properties are all derivable as a consequence of two physical assumptions implicit in those definitions, namely, the Maxwell-Lorentz theory of electromagnetism and the postulates of causality, linearity, time-invariance, and passivity. These two assumptions alone thus constitute an axiomatic physical foundation for electromagnetic system theory.

TABLE OF CONTENTS

Introduction.....	1
Nature of the physical assumptions.....	2
Temporal description of systems.....	6
Spectral description of systems.....	16
Spectral description of signals.....	24
Exact response to a sinusoidal input signal.....	28
Conclusion.....	32

1. INTRODUCTION

The physical meaning of the definitions on which electromagnetic system theory is based is not always clear. The spectral description of a system, for example, is frequently defined by its response to a time-harmonic signal, and yet true time-harmonic signals are non-physical because they would have had to exist forever. The spectral description of a signal as the Fourier transform of its complete time history, including the future as well as the past, is another questionable definition because of the implication that the future can influence the present. The trouble here will be seen to lie in the fact that the concept of signal spectrum has no physical meaning when divorced from some system on which it must act to be observed. This illustrates clearly that a thorough re-examination of the physical foundation of system theory could be worthwhile. Instead of basing the theory on a collection of abstract definitions it would be preferable to base it on verifiable physical assumptions. Then if any physical inconsistencies are found to exist they can be traced back directly to some fault in the assumptions. That is the program of the present paper. A general theory of two-port systems--circuits, microwave networks, antennas, etc.--will be developed here from just two physical assumptions. One is that the Maxwell-Lorentz theory of electromagnetism, which has been verified in so many different ways that it is frequently considered to be a law of nature, is valid. The other is the usual set of postulates on system behavior: causality, linearity, time-invariance, and passivity. From these two assumptions alone all of the properties usually attributed to electromagnetic systems will be seen to follow. In the case of lumped constant circuits these assumptions provide a physical foundation for circuit theory that is not possible to establish from circuit theory itself.

2. NATURE OF THE PHYSICAL ASSUMPTIONS

a) The Maxwell-Lorentz theory of electromagnetism.

It is a fact of experience that all known electromagnetic phenomena are described by the Maxwell-Lorentz theory. Hence the assumption of its validity is an almost mandatory starting point for any development of electromagnetic system theory.

Of Maxwell's four equations his two curl equations

$$\nabla \times \underline{\underline{E}} = - \frac{\partial \underline{\underline{B}}}{\partial t} \quad , \quad (1)$$

$$\nabla \times \underline{\underline{H}} = \underline{\underline{j}} + \frac{\partial \underline{\underline{D}}}{\partial t} \quad , \quad (2)$$

are the only two that are independent (the two divergence equations are automatically satisfied by them for any physical field, which must be zero prior to creation of the source at some finite time in the past). But they provide only two relationships between the five field functions $\underline{\underline{E}}$, $\underline{\underline{B}}$, $\underline{\underline{A}}$, $\underline{\underline{j}}$, $\underline{\underline{H}}$, which means that three more are needed. The other three are provided by Lorentz's electron theory for the physical properties of the electromagnetic medium of which the system is constituted. These properties, arising in part from the properties of space and the rest from the internal electronic structure of the material medium, determine the dependence of $\underline{\underline{j}}$ and $\underline{\underline{D}}$ on the force field $\underline{\underline{E}}$ and of $\underline{\underline{H}}$ on the force field $\underline{\underline{B}}$. Once these constitutive relationships have been determined at every point in the system the fields $\underline{\underline{E}}$ and $\underline{\underline{B}}$ everywhere can be determined uniquely from (1) and (2) by imposing boundary conditions.

As a specific illustration consider the electromagnetic system shown in Fig. 1. Physical input signals are always real functions of time, taken here to be a voltage $v(t')$ applied at time t' at the reference location $(0,0,0)$ to a coaxial cable feeding a horn antenna. This antenna, by itself, constitutes a one-port system. At every point (x,y,z) in that system the input signal will produce five field functions at time t .

They will be determined everywhere by the time history of the input voltage and by the distribution of matter in and about the horn. If a second antenna is present, such as the receiving dipole and reflector shown at the right, then the fields everywhere will be determined by the distribution of matter in and about that antenna, too. But usually we are not interested in the fields everywhere. In most cases the interest centers either on $\underline{V}(0,0,0,t)$ at the input port, from which the input current $i(t)$ to the system can be determined, or on $\underline{E}(x,y,z,t)$ or $\underline{V}(x,y,z,t)$ at the output port (x,y,z) from which the output voltage $v(t)$ or output current $i(t)$, respectively, can be determined. In this sense the two-port antenna system illustrated here is no different from a circuit or microwave network. It is for this reason that the theory to be developed here is believed to be quite general.

A question arises regarding synchronization of the clocks by which time t' at the input port and time t at the observation point (x,y,z) are to be measured. Einstein [1] answered this question when he pointed out that similar clocks at rest in a common frame of reference can be synchronized by means of light signals. A common time can then be assigned to an event taking place anywhere in that reference frame. But no common time exists for clocks in relative motion. Thus the results of the system theory to be developed here are limited to points of observation that are at rest with respect to a frame fixed in the system.

b) Postulates on system behavior.

The second and final assumption is that the system is causal, linear, time-invariant, and passive. None of these properties is indigenous to the Maxwell-Lorentz theory as a whole, although some are in part. Thus they represent an independent assumption. An important consequence of this assumption is that it leads to a description of the medium in

terms of the complex spectral functions $\epsilon(x, y, z, \omega)$, $\sigma(x, y, z, \omega)$, and $\mu(x, y, z, \omega)$ whose properties are known from experience to represent a wide range of physically important media.

Causality - The familiar one-way flow of events from cause to effect is so natural to our senses that it hardly seems necessary to postulate it. Surely there must be some law that asserts the necessity of causality in nature. But the surprising fact of the matter seems to be that none of the known laws of physics show any distinction between past and future. On a microscopic scale they are all completely reversible in time. This includes Maxwell's equations, which admit of advanced solutions just as readily as retarded solutions, even though no macroscopic evidence of advanced solutions has ever been found. The response of \underline{D} to \underline{E} and of \underline{H} to \underline{B} at every point in a vacuum is clearly reversible, too, their relationship being a simple proportionality. Presumably the response of each atom in a material medium is also reversible. They why do electromagnetic phenomena on a macroscopic scale appear to be irreversible? The answer is that they may actually be reversible, in principle, but that the probability of physical phenomena reversing their sequence of events in time is so small as to be practically zero [2]. On this basis causality on a macroscopic scale should not be regarded as being an absolute law of nature but rather as being a chance event that is so highly probable as to be virtually certain to occur. As a practical matter, then, the postulate of causality can be considered to be a necessary feature of any system theory. It was long ago incorporated into circuit theory as a necessary requirement for physical realizability.

Linearity - Any system for which a linear combination of exciting signals produces the same linear combination of responses is said to be linear. Maxwell's equations are linear (any linear combination of solutions is also a solution) and free space is linear (\underline{D} is proportional to \underline{E} , and \underline{H} to \underline{B}),

but material media are never exactly linear. Hence the postulate of linearity is really an assumption on the medium alone. A wide range of media of practical interest are nearly linear; Δ and \downarrow found by experiment in passive systems are frequently related almost linearly to \underline{E} alone, and \underline{X} to \underline{Q} alone.

Although the postulate of linearity is never satisfied exactly in material media, and may not even be a reasonable approximation in some media, it nevertheless is believed to be an absolutely indispensable feature of any theory of electromagnetic systems. Without it the influence of the input signal on the fields in the system could never be separated from the influence of the system itself. As a consequence the field equations would have to be re-solved completely at every moment in time for ever new input signal. Only for linear systems can the fields of an arbitrary input signal be obtained by solving the field equations once and for all for the fields of some elementary input signal from which the fields of an arbitrary input signal can be constructed. The importance of being able to separate the description of the system from that of its input signal can hardly be overstated, as it is this that makes system theory possible.

Time-invariance - Any system whose properties remain fixed in time is said to be time-invariant. Maxwell's equations are time-invariant (they are unchanged by a shift in the time reference) and so is free space. Hence the postulate of time-invariance, like that of linearity above, is an assumption solely on the material medium alone. As in the case of linearity it is a good approximation for many physical media of practical interest.

Unlike the postulates of causality and linearity, neither time-invariance nor the postulate of passivity to follow appear to be necessary ingredients of a theory of systems. They are included here because they are implicit in the system

properties usually assumed and because they provide a basis for the derivation of those properties.

Passivity - A system that contains no internal energy sources is said to be passive. Free space and material media are both passive properties of a system because they introduce no new sources of energy. Internal energy sources are represented solely by an active term in the current density \mathbf{j} in Maxwell's equations. Hence the postulate of passivity is an assumption on Maxwell's equations alone.

Those properties of the above four postulates that are implicit in the Maxwell-Lorentz theory of electromagnetism are summarized in Table 1.

TABLE 1

	Field Equations	Free Space	Material Media
Causality			
Linearity	X	X	
Time-invariance	X	X	
Passivity		X	X

For the others the postulates represent an additional assumption. In contrast to the universal truth that the Maxwell-Lorentz theory appears to represent, these additional assumptions may be only approximately true.

3. TEMPORAL DESCRIPTION OF SYSTEMS

Any system that satisfies the Maxwell-Lorentz theory can be characterized completely by a set of five real field functions produced everywhere in the system by a vanishingly short pulse of fixed area (an impulse) at its input port. Two of

these fields, \underline{E} and \underline{B} , produce the Lorentz force on electric charge. They are related to each other directly by Faraday's law (1) and indirectly, through the constitutive properties of the medium, by Maxwell's modification (2) of Ampere's law. They act locally as exciting forces on the electronic structure of the medium to produce \underline{D} , \underline{J} , and \underline{H} that represent the other three fields. The objective here is to describe the macroscopic properties of these five impulse response field functions. From these impulse response fields all of the other results of system theory will be derived.

The fact that a vanishingly short pulse is the characteristic input signal by which all input signals are constructed physically is implicit from the way in which the temporal record of signals develop (Fig. 2). From the moment $t'=0$ at which the signal record first begins it develops as an ever lengthening continuum of pulses of infinitesimal area (height times duration) up to time $t'=t$, the present moment "now", beyond which it is zero because the future signal history has not yet occurred. The characteristic temporal description of a system is thus the field responses to a unit impulse (a vanishingly short pulse of unit area). The response to an arbitrary input signal can then be constructed mathematically by superposition of the responses to each infinitesimal element of its area. But this raises two questions. One is that a vanishingly short pulse of unit area must have infinite height and hence could never exist physically, not to mention the fact that no physical system could survive such a shock. The other is the implicit assumption that the response of an impulse is proportional to its area. A precise answer to these and other related questions will now be given before proceeding further.

The second question will be answered first by proving, perhaps for the first time, that the response to an impulse is, indeed, proportional to its area. It is a direct consequence

of the postulates of linearity and time-invariance in the limit of vanishing pulse duration.

Theorem: The responses of an electromagnetic system to a vanishingly short pulse are proportional to the area of that pulse.

Proof: Let a rectangular pulse of given area and non-zero duration Δt be impressed upon a system free of initial excitation. After the pulse ceases each response function must originate solely from energy stored within the (passive) system. Each response everywhere in the system will be a definite, unique function of time that characterizes that particular system. To prove that it becomes proportional to area as the pulse duration vanishes it is sufficient to show that it becomes proportional to height and to duration separately. That it is proportional to height is obvious from linearity, regardless of duration; i.e., changing the pulse height by any constant factor will change the response everywhere by exactly that same factor. Proportionality to duration, however, is not at all obvious. To establish it imagine the pulse sliced into n equal time intervals $\Delta t/n$. This is indicated in Fig. 3, along with an example of a hypothetical current response $i(t)$ at the input port to illustrate one particular response function of interest. The response due to each of the identical slices is the same for all, because of time-invariance, but each is displaced from its immediate neighbors by the time interval $\Delta t/n$. Furthermore the sum of these identical responses must add up to exactly the response of the original pulse because of linearity. If it weren't for the time displacement between responses, then, the response to each of the n slices would be exactly $1/n$ of that of the original pulse; i.e., the response of each slice would be proportional to its duration. But when Δt vanishes the time displacement between responses vanishes also. Thus the response does, in fact, become proportional to pulse duration as the duration vanishes. It is concluded, therefore, that

the response to a vanishingly short pulse is proportional everywhere in space and in time to its height and to its duration separately, and hence to its area.

Two questions arise regarding the current response right at the input port. One is the behavior of that response within the duration of the impulse itself, and the other is the finiteness of the energy supplied to the system by the impulse. The two are related and can be answered together. Note first that the current response to an input voltage pulse of finite height cannot jump discontinuously, because of inductance that will always be present in any physical system. It must change continuously from its initial value of zero. Thus the unknown response current within the duration of the pulse can be expressed as a power series that begins with the linear term,

$$i(t) = \frac{A}{\Delta t} (a_1 t + a_2 t^2 + \dots) , \quad 0 \leq t \leq \Delta t . \quad (3)$$

By showing the dependence on pulse height explicitly in (3) through the proportionality factor $A/\Delta t$, where the pulse area is denoted by A (1 volt-sec), the expansion coefficients a_n are independent of both time and the pulse height. As the duration Δt vanishes, then, it is evident from (3) that the current response must become linear unless the leading coefficient a_1 is precisely zero. To determine whether it could actually be zero, look at the expression for energy supplied to the system by the voltage pulse $v(t) = A/\Delta t$, $0 < t \leq \Delta t$:

$$\int_0^{\Delta t} v(t) i(t) dt = \frac{A^2 a_1}{2} + \frac{A^2 a_2}{3} \Delta t + \frac{A^2 a_3}{4} (\Delta t)^2 + \dots \quad (4)$$

If a_1 were zero then the energy supplied by a unit impulse ($\Delta t \rightarrow 0$) would be zero, which would mean zero response thereafter. It is concluded that a_1 cannot be zero. The current response within the vanishing duration of the pulse must then

become linear. At each moment within the pulse it becomes proportional to the area that exists at that moment, which means that the theorem above applies even within the duration of the impulse itself. In the limit as Δt approaches zero, then, the current response (3) within the duration of the impulse approaches an abrupt, but linear, change from its initial value of zero to the finite value $A\alpha_0$ at the end, as indicated in Fig. 4. And the energy (4) supplied to the system always remains finite as $\Delta t \rightarrow 0$, never becoming infinite as is sometimes suggested by such non-physical systems as a pure resistance. These are general characteristics that are true for all physical systems satisfying the assumptions made earlier. They are not necessarily true for non-physical systems, however, such as a lumped constant circuit without inductance at its input.

Returning to the original question, the fact that a pulse of infinite height could never be produced in practice can now be seen to be of no concern. According to the theorem it is only the area of a pulse of vanishing length that is important. To determine impulse response experimentally it is sufficient that the measured response to a pulse of fixed height and decreasing duration reach its limiting functional form before sinking into the background noise. This limiting response function will then be proportional to the response to a unit impulse. The proportionality factor needed to normalize it to that of a unit impulse is just the area of the pulse.

With the above clarification of the physical meaning of impulse response it is now clear why the characteristic temporal description of electromagnetic systems is their field responses to an impulse. But the physical nature of these field responses is quite different for \underline{E} and \underline{B} than for \underline{D} , \underline{J} , and \underline{M} . The \underline{E} and \underline{B} fields produce the Lorentz force, hence are the only exciting force fields (causes) in the medium. They act locally to produce \underline{D} , \underline{J} , and \underline{M} (effects) at each

point. In the case of material media this causal relationship is quite clear experimentally for the induced current density \underline{j} of free electrons and for the polarization part of \underline{A} and of \underline{H} . For the vacuum part, however, there appears to be no unique causal relationship because the vacuum part of \underline{A} is just proportional to \underline{E} and that of \underline{H} to \underline{B} , with no indication of which is the cause and which the effect. In the absence of further knowledge the causal relationship for the vacuum part will be considered here to be the same as that of the polarization part. Thus \underline{A} , \underline{j} , and \underline{H} will be considered to be totally subservient to \underline{E} and \underline{B} , being only local responses of the medium that are produced by the global exciting forces \underline{E} and \underline{B} of the system. The term "global" is used here to indicate dependence of a response upon the system as a whole at a point some distance away from its excitation, in contrast to the term "local" used to indicate dependence upon the medium alone in the immediate neighborhood of its point of excitation.

In view of this essential physical difference between the global forces \underline{E} and \underline{B} and their local responses \underline{A} , \underline{j} , and \underline{H} , the five impulse response functions characterizing the system will be taken here to be the \underline{E} and \underline{B} responses of the system to a unit impulse of voltage at the input port and the \underline{A} , \underline{j} , and \underline{H} responses of the medium to a unit impulse of \underline{E} and \underline{B} at that same point in the medium. The theorem above on response of a system applies equally well to these three local responses of the material part of the medium because they, too, originate from stored energy. In this case it is just the energy stored locally in the electronic structure of the material medium.

The two global impulse response functions will be denoted here by $\hat{\underline{E}}(x, y, z, t-t')$ and $\hat{\underline{B}}(x, y, z, t-t')$ representing the \underline{E} and \underline{B} field responses at the point (x, y, z) at time t to a unit impulse of voltage at the input terminals $(0,0,0)$ at time t' . The fact that their time dependence involves only the difference $t-t'$ is a consequence of the postulate of time-invariance. The three

local impulse response functions will be denoted by $\hat{\epsilon}(x, y, z, t-t')$, $\hat{\sigma}(x, y, z, t-t')$, and $\hat{\mu}^{-1}(x, y, z, t-t')$, the first two representing the \underline{A} and \underline{J} field responses, respectively, at the point (x, y, z) at time t to a unit impulse of \underline{E} at that same point at time t' , with the third representing the \underline{H} field response to a unit impulse of \underline{B} . These local functions describe the instantaneous constitutive properties of the medium alone, representing instantaneous dielectric susceptibility, conductivity, and inverse magnetic permeability, respectively. They will be seen later from (26) to be responsible for the dispersive properties of the medium. They cannot be measured directly but only deduced from their effect on the global response functions.

The global response functions \underline{E} and \underline{B} will always be zero prior to the earliest moment $t=t'+\frac{r}{c}$ at which the field of an impulse could arrive at the point (x, y, z) from the input port:

$$\begin{aligned}\hat{\underline{E}}(x, y, z, t-t') &= 0 \\ \hat{\underline{B}}(x, y, z, t-t') &= 0, \quad t \leq t' + \frac{r}{c}.\end{aligned}\tag{5}$$

This combined statement of causality and propagation delay includes the moment $t=t'+\frac{r}{c}$ of arrival of the leading edge of the impulse because the area of the retarded pulse at that moment is still zero. The local response functions $\hat{\epsilon}$, $\hat{\sigma}$, and $\hat{\mu}^{-1}$ experience no propagation delay, so

$$\begin{aligned}\hat{\epsilon}(x, y, z, t-t') &= 0 \\ \hat{\sigma}(x, y, z, t-t') &= 0 \\ \hat{\mu}^{-1}(x, y, z, t-t') &= 0, \quad t \leq t'.\end{aligned}\tag{6}$$

The \underline{E} and \underline{B} fields produced by an arbitrary input voltage $v(t')$ can now be obtained by summing their responses produced by all of the elements of area $v(t')\Delta t'$, as illustrated in Fig. 5,

$$\begin{aligned}\underline{\Xi}(x, y, z, t) &= \int_0^{t-\frac{\Delta}{c}} v(t') \hat{\underline{\Xi}}(x, y, z, t-t') dt' , \\ \underline{\Theta}(x, y, z, t) &= \int_0^{t-\frac{\Delta}{c}} v(t') \hat{\underline{\Theta}}(x, y, z, t-t') dt' .\end{aligned}\quad (7)$$

The summation extends from time $t'=0$, at which the signal begins, to the time $t'=t-\frac{\Delta}{c}$, beyond which the impulse responses are zero by (5). The $\underline{\Delta}$ and $\underline{\downarrow}$ fields produced locally by $\underline{\Xi}$, and the $\underline{\chi}$ field produced locally by $\underline{\Theta}$, are obtained in the same way by summing the local responses of the medium,

$$\begin{aligned}\underline{\Delta}(x, y, z, t) &= \int_{\chi_k}^t \underline{\Xi}(x, y, z, t') \hat{\underline{\epsilon}}(x, y, z, t-t') dt' , \\ \underline{\downarrow}(x, y, z, t) &= \int_{\chi_c}^t \underline{\Xi}(x, y, z, t') \hat{\underline{\sigma}}(x, y, z, t-t') dt' , \\ \underline{\chi}(x, y, z, t) &= \int_{\chi_c}^t \underline{\Theta}(x, y, z, t') \hat{\underline{\mu}}^{-1}(x, y, z, t-t') dt' .\end{aligned}\quad (8)$$

The summation here extends from the time $t'=\frac{\Delta}{c}$ at which the exciting forces $\underline{\Xi}$ and $\underline{\Theta}$ first appear to the time t beyond which the response functions of the medium are zero by (6). These constitutive relationships (8) between the temporal field functions are quite different from the simple proportionalities that will be seen later to connect the corresponding spectral functions. The integrals (7) and (8) are exact representations for the fields everywhere in terms of the five temporal functions $\underline{\Xi}$, $\underline{\Theta}$, $\hat{\underline{\epsilon}}$, $\hat{\underline{\sigma}}$, and $\hat{\underline{\mu}}^{-1}$ that characterize the system completely.

The exciting forces $\underline{\Xi}$ and $\underline{\Theta}$ are produced by the input voltage $v(t')$ by (7) and they, in turn, produce $\underline{\Delta}$, $\underline{\downarrow}$, and $\underline{\chi}$ as in (8). Hence they can be eliminated from (8). Putting $\underline{\Xi}$ and $\underline{\Theta}$ from (7) (after extending the upper limit from $t-\frac{\Delta}{c}$ to infinity as required by (5) for a complete description of $\underline{\Xi}$ and $\underline{\Theta}$) into (8) and interchanging the order of integration a

simple calculation shows that (8) becomes

$$\begin{aligned}\underline{D}(x, y, z, t) &= \int_0^{t-\frac{a}{c}} v(t') \underline{\hat{D}}(x, y, z, t-t') dt' , \\ \underline{f}(x, y, z, t) &= \int_0^{t-\frac{a}{c}} v(t') \underline{\hat{f}}(x, y, z, t-t') dt' , \\ \underline{h}(x, y, z, t) &= \int_0^{t-\frac{a}{c}} v(t') \underline{\hat{h}}(x, y, z, t-t') dt' ,\end{aligned}\tag{9}$$

where the new functions $\underline{\hat{D}}$, $\underline{\hat{f}}$, and $\underline{\hat{h}}$ represent the following integrals,

$$\begin{aligned}\underline{\hat{D}}(x, y, z, t-t') &\triangleq \int_{a/c}^{t-t'} \underline{\hat{e}}(x, y, z, t-t'-t'') \underline{\hat{E}}(x, y, z, t'') dt'' , \\ \underline{\hat{f}}(x, y, z, t-t') &\triangleq \int_{a/c}^{t-t'} \underline{\hat{\sigma}}(x, y, z, t-t'-t'') \underline{\hat{E}}(x, y, z, t'') dt'' , \\ \underline{\hat{h}}(x, y, z, t-t') &\triangleq \int_{a/c}^{t-t'} \underline{\hat{\mu}}^{-1}(x, y, z, t-t'-t'') \underline{\hat{B}}(x, y, z, t'') dt'' .\end{aligned}\tag{10}$$

From (6) it is evident that these new functions (10) must satisfy the conditions

$$\begin{aligned}\underline{\hat{D}}(x, y, z, t-t') &= 0 , \\ \underline{\hat{f}}(x, y, z, t-t') &= 0 , \\ \underline{\hat{h}}(x, y, z, t-t') &= 0 , \quad t \leq t' + \frac{a}{c} ,\end{aligned}\tag{11}$$

which correspond to the conditions (5) on $\underline{\hat{E}}$ and $\underline{\hat{B}}$. Physically they represent the local response of the medium to the global exciting forces $\underline{\hat{E}}$ and $\underline{\hat{B}}$. Since they have a global dependence through $\underline{\hat{E}}$ and $\underline{\hat{B}}$ they represent a global alternative to the purely local functions $\underline{\hat{e}}$, $\underline{\hat{\sigma}}$, and $\underline{\hat{\mu}}^{-1}$ characterizing the medium alone.

In summary, a complete temporal description of the system is provided by: (a) the global impulse response functions $\hat{\underline{E}}$ and $\hat{\underline{B}}$, and (b) either the global impulse response functions $\hat{\underline{A}}$, $\hat{\underline{J}}$, and $\hat{\underline{X}}$ characterizing the system as a whole or the local impulse response functions $\hat{\underline{E}}$, $\hat{\underline{\sigma}}$, and $\hat{\underline{\mu}}$ characterizing the medium alone.

The admittance functions of circuit theory come directly from these impulse response fields. The input admittance, for example, comes from the current response produced at the input port by a unit impulse of voltage. For a perfectly conducting coaxial input the current response to any arbitrary driving voltage is obtained from the line integral of the tangential component of the \underline{X} field (9) around the center conductor at the input port,

$$\begin{aligned} i(t) &= \oint \underline{X}(0,0,0,t) \cdot d\underline{\ell} \\ &= \int_0^t v(t') \left[\oint \hat{\underline{X}}(0,0,0,t-t') \cdot d\underline{\ell} \right] dt' . \end{aligned} \quad (12)$$

From this it is evident that the current response $\hat{y}(t-t')$ at the input at time t due to a unit impulse of voltage at time t' is produced by $\hat{\underline{X}}(0,0,0,t-t')$,

$$\hat{y}(t-t') = \oint \hat{\underline{X}}(0,0,0,t-t') \cdot d\underline{\ell} . \quad (13)$$

All of the preceding properties are exact and quite general. But another property -- a bound on the temporal behavior of the impulse response field functions -- will be needed for which the result to be established may be only approximate, namely, that the impulse response fields must remain within an exponentially decaying envelope. This follows from the finiteness of the input energy (4), which becomes stored in the

fields of the passive system and becomes the only source of energy available after the input impulse ceases. This stored energy decreases with time because of dissipation present in all physical systems. Over any given interval of time it will decrease to some fixed fraction, on the average, of the value that it had at the beginning of that interval. Thus the total energy in the system must remain within an exponentially decaying envelope. By making the reasonable but approximate assumptions that the energy densities at every point will also have this same behavior and that the energy densities are proportional to $\hat{\underline{E}} \cdot \hat{\underline{E}}$ and $\hat{\underline{H}} \cdot \hat{\underline{H}}$ (exactly true for non-dispersive media only) it follows that the impulse response fields themselves must remain within an exponential envelope. For $\hat{\underline{E}}$, for example,

$$|\hat{\underline{E}}(x, y, z, \tau)| < A e^{-a\tau}, \quad \tau > \frac{1}{2}, \quad (14)$$

where $A(x, y, z)$ and $a(x, y, z)$ are real functions independent of τ .

4. SPECTRAL DESCRIPTION OF SYSTEMS

The well known fact that the spectral functions $\underline{E}(x, y, z, \omega)$, $\underline{B}(x, y, z, \omega)$, etc. in time-harmonic system theory are the Fourier integral of their corresponding impulse response functions will be shown here to be a necessary consequence of solving Maxwell's equations by separation of variables. The complex spectral variable ω will enter as the separation parameter, and the spectral functions as the spatial factor, in the separated solution. This derivation of the Fourier relationship between the temporal and spectral description of systems establishes clearly the fact that these spectral functions are strictly a mathematical description of the system alone, with their physical origin residing solely in the impulse response functions from which they arose. Recognition of this essential fact should strip away much of the prevailing confusion on the

physical meaning of the spectral functions that has arisen from mistaking the consequence that they have a physical interpretation as steady-state complex amplitudes for real positive values of ω as being, instead, their physical origin. It also gives a deeper meaning to the impedance and transfer functions of circuit theory than is possible from circuit theory itself.

The five impulse response field functions describe physical fields, even though they have dimensions of field per volt-second. Hence they must satisfy Maxwell's equations everywhere in the system,

$$\begin{aligned}\nabla \times \hat{\underline{E}} &= - \frac{\partial \hat{\underline{B}}}{\partial t} \\ \nabla \times \hat{\underline{H}} &= \hat{\underline{J}} + \frac{\partial \hat{\underline{D}}}{\partial t}\end{aligned}\quad (15)$$

The corresponding spectral functions now come from separation of the time variable from the spatial variables in (15). The usual procedure of first reducing (15) to a wave equation is invalid for material media because of the effects of dispersion. The general solution can be obtained only by separating variables in the two field equations (15) directly. Surprisingly, this seems not to have been done before.

An elementary solution for $\hat{\underline{E}}(x, y, z, \tau)$ will be sought in the form of the product $\underline{E}(x, y, z)T(\tau)$, that for $\hat{\underline{B}}(x, y, z, \tau)$ in the form of the product $\underline{B}(x, y, z)T(\tau)$, etc., with a single time function common to all five fields. Then from (15),

$$\begin{aligned}(\nabla \times \underline{E})_{\lambda} T &= -B_{\lambda} T' \\ (\nabla \times \underline{H})_{\lambda} T &= J_{\lambda} T + D_{\lambda} T'\end{aligned}\quad (16)$$

where λ denotes each of the three rectangular coordinates and T' denotes the derivative with respect to τ . Dividing the

first through by $B_\lambda T$ and the second by $D_\lambda T$,

$$\begin{aligned} \frac{(\nabla \times \underline{E})_\lambda}{B_\lambda} &= -\frac{T'}{T} \\ \frac{(\nabla \times \underline{H})_\lambda - J_\lambda}{D_\lambda} &= \frac{T'}{T} \end{aligned} \quad (17)$$

The left side of each is a function of the space coordinates only while the right side is a function of time only, hence each must be independent of both space and time. Denoting the value of the first by an arbitrary complex number* $-i\omega$, the value of the second must be $+i\omega$. For each value of ω it is concluded from (17) that the spacial functions must satisfy

$$\begin{aligned} \nabla \times \underline{E}(x, y, z, \omega) &= -i\omega \underline{B}(x, y, z, \omega) \\ \nabla \times \underline{H}(x, y, z, \omega) &= \underline{J}(x, y, z, \omega) + i\omega \underline{D}(x, y, z, \omega) \end{aligned} \quad (18)$$

and that the time function must satisfy

$$\frac{\partial T(\tau, \omega)}{\partial \tau} - i\omega T(\tau, \omega) = 0 \quad (19)$$

Solving (19) for T ,

$$T(\tau, \omega) = C(\omega) e^{i\omega\tau} \quad (20)$$

* This choice of notation is solely a matter of convenience in order that the end results conform with the conventional spectral notation. If the convention $e^{-i\omega\tau}$ is preferred instead of $e^{+i\omega\tau}$ in the result (22), a choice sometimes seen in the physics literature, it is simply a matter of choosing $+i\omega$ for the complex number here rather than $-i\omega$.

In view of the linearity of Maxwell's equations the most general solution to (15) obtainable by this process of separating variables is a linear superposition of these elementary solutions,

$$\begin{aligned}
 \hat{\underline{E}}(x, y, z, \tau) &= \frac{1}{2\pi} \int_{\underline{C}_1} \underline{E}(x, y, z, \omega) e^{i\omega\tau} d\omega, \\
 \hat{\underline{B}}(x, y, z, \tau) &= \frac{1}{2\pi} \int_{\underline{C}_2} \underline{B}(x, y, z, \omega) e^{i\omega\tau} d\omega, \\
 \hat{\underline{D}}(x, y, z, \tau) &= \frac{1}{2\pi} \int_{\underline{C}_3} \underline{D}(x, y, z, \omega) e^{i\omega\tau} d\omega, \\
 \hat{\underline{J}}(x, y, z, \tau) &= \frac{1}{2\pi} \int_{\underline{C}_4} \underline{J}(x, y, z, \omega) e^{i\omega\tau} d\omega, \\
 \hat{\underline{H}}(x, y, z, \tau) &= \frac{1}{2\pi} \int_{\underline{C}_5} \underline{H}(x, y, z, \omega) e^{i\omega\tau} d\omega,
 \end{aligned} \tag{21}$$

where the spatial functions \underline{E} , \underline{B} , \underline{D} , \underline{J} , \underline{H} must satisfy (18) but are otherwise undetermined (the factor $2\pi\mathcal{L}(\omega)$ has been absorbed into each) and where the integration contours \underline{C}_k in the complex ω -plane are also undetermined.

The only remaining problem is to determine the spatial functions and integration contours such that each of the integrals (21) will be an exact representation for the corresponding impulse response function. It is a problem that always occurs whenever partial differential equations are solved by the method of separation of variables because the method itself does not guarantee completeness of the solution. The answer lies in Plancherel's proof [3] of the Fourier integral theorem, which states that if the integration contours are restricted to the real axis in the complex ω -plane the integral representations

$$\begin{aligned}
\hat{\underline{E}}(x, y, z, \tau) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \underline{E}(x, y, z, \omega) e^{i\omega\tau} d\omega, \\
\hat{\underline{B}}(x, y, z, \tau) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \underline{B}(x, y, z, \omega) e^{i\omega\tau} d\omega, \\
\hat{\underline{D}}(x, y, z, \tau) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \underline{D}(x, y, z, \omega) e^{i\omega\tau} d\omega, \\
\hat{\underline{J}}(x, y, z, \tau) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \underline{J}(x, y, z, \omega) e^{i\omega\tau} d\omega, \\
\hat{\underline{H}}(x, y, z, \tau) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \underline{H}(x, y, z, \omega) e^{i\omega\tau} d\omega,
\end{aligned} \tag{22}$$

for all square integrable functions on $-\infty < \tau < \infty$ will each be complete in the sense of zero mean-square error provided only that the spectral functions are chosen to be

$$\begin{aligned}
\underline{E}(x, y, z, \omega) &\triangleq \int_{r/c}^{\infty} \hat{\underline{E}}(x, y, z, \tau) e^{-i\omega\tau} d\tau, \\
\underline{B}(x, y, z, \omega) &\triangleq \int_{r/c}^{\infty} \hat{\underline{B}}(x, y, z, \tau) e^{-i\omega\tau} d\tau, \\
\underline{D}(x, y, z, \omega) &\triangleq \int_{r/c}^{\infty} \hat{\underline{D}}(x, y, z, \tau) e^{-i\omega\tau} d\tau, \\
\underline{J}(x, y, z, \omega) &\triangleq \int_{r/c}^{\infty} \hat{\underline{J}}(x, y, z, \tau) e^{-i\omega\tau} d\tau, \\
\underline{H}(x, y, z, \omega) &\triangleq \int_{r/c}^{\infty} \hat{\underline{H}}(x, y, z, \tau) e^{-i\omega\tau} d\tau.
\end{aligned} \tag{23}$$

The lower limit is shown as r/c because of causality. The sole condition here on the impulse response functions is that they be square integrable. It is a condition that is believed to be satisfied for all physical systems because of the approximate result (14) that impulse responses must remain within an exponentially decaying envelope. Thus the five spectral functions describing the system must be of the form (23) and must satisfy the partial differential equations (18) usually referred to as time-harmonic field equations.

The Fourier integral relationship (23) between the impulse response functions and their corresponding spectral functions arises as a consequence of separating the temporal from the spatial dependence in the solution of Maxwell's equations. It is of interest to note that this is the only point at which Fourier theory enters into the present development of electromagnetic system theory.

To achieve the important property of completeness in the integral representation (22) the integration contours were restricted to the real axis of the complex ω -plane. But the contours can be deformed at will within any region of analyticity of the integrand without changing the value of the integrals. The complex spectral functions defined by the integral representation (23) each have real and imaginary parts that satisfy the Cauchy-Riemann conditions, as can be seen by direct calculation, which means that the region of analyticity in the complex ω -plane is at least that over which the integrals converge. They will always converge wherever the imaginary part of ω is less than some real positive number α that describes the rate of decay (14) of the exponential bound on the impulse response functions. Hence each of the spectral functions in (22) is assured of being analytic in at least the lower half plane $\omega_i < \alpha$. Except at its singular points each must also be analytic everywhere in the upper half plane, too, even in regions where the integral representation (23) is no longer valid. The analytic character of a function of a complex variable is determined not by the region of validity of a particular representation but solely by the nature of its singularities. The integrand of each of the integrals (22) will then be analytic everywhere in at least the lower half plane $\omega_i < \alpha$, since the factor $e^{i\omega t}$ is an entire function of ω , which means that the integration contours can each be deformed to lie anywhere within that region without in any way changing the value of the integral. Thus there is a great deal of

freedom in the choice of the contours C_k in the original solution (21).

The spectral description of the medium alone can now be obtained in terms of the local impulse response functions $\hat{\epsilon}$, $\hat{\sigma}$, $\hat{\mu}^{-1}$ by expressing $\underline{\hat{\epsilon}}$ and $\underline{\hat{\theta}}$ in (10) by (22) and interchanging the order of integration,

$$\begin{aligned}\underline{\hat{\Delta}}(x, y, z, \tau) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\tau} \hat{\epsilon}(x, y, z, \tau - t') e^{i\omega t'} dt' \right] \underline{E}(x, y, z, \omega) d\omega, \\ \underline{\hat{J}}(x, y, z, \tau) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\tau} \hat{\sigma}(x, y, z, \tau - t') e^{i\omega t'} dt' \right] \underline{E}(x, y, z, \omega) d\omega, \\ \underline{\hat{X}}(x, y, z, \tau) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\tau} \hat{\mu}^{-1}(x, y, z, \tau - t') e^{i\omega t'} dt' \right] \underline{B}(x, y, z, \omega) d\omega.\end{aligned}\tag{24}$$

The lower limit on t' was extended from $\frac{\tau}{c}$ to $-\infty$ as required for the full description of $\underline{\hat{\epsilon}}$ and $\underline{\hat{\theta}}$. After a change of variable $t' = \tau - \tau'$ these reduce to

$$\begin{aligned}\underline{\hat{\Delta}}(x, y, z, \tau) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \epsilon(x, y, z, \omega) \underline{E}(x, y, z, \omega) e^{i\omega\tau} d\omega, \\ \underline{\hat{J}}(x, y, z, \tau) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \sigma(x, y, z, \omega) \underline{E}(x, y, z, \omega) e^{i\omega\tau} d\omega, \\ \underline{\hat{X}}(x, y, z, \tau) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{\mu}^{-1}(x, y, z, \omega) \underline{B}(x, y, z, \omega) e^{i\omega\tau} d\omega,\end{aligned}\tag{25}$$

where the spectral functions $\epsilon(x, y, z, \omega)$, $\sigma(x, y, z, \omega)$, $\hat{\mu}^{-1}(x, y, z, \omega)$ describing the medium alone represent the integrals

$$\begin{aligned}
\epsilon(x, y, z, \omega) &\triangleq \int_0^{\infty} \hat{\epsilon}(x, y, z, \tau) e^{-i\omega\tau} d\tau, \\
\sigma(x, y, z, \omega) &\triangleq \int_0^{\infty} \hat{\sigma}(x, y, z, \tau) e^{-i\omega\tau} d\tau, \\
\mu^{-1}(x, y, z, \omega) &\triangleq \int_0^{\infty} \hat{\mu}^{-1}(x, y, z, \tau) e^{-i\omega\tau} d\tau.
\end{aligned}
\tag{26}$$

The latter are the familiar functions usually associated with the properties of dielectric susceptibility, conductivity, and the reciprocal of magnetic permeability. They describe the dispersive properties of the medium. It is now clear why $\hat{\epsilon}$, $\hat{\sigma}$, $\hat{\mu}^{-1}$ represent the temporal description of those same properties.

Comparing (25) with (22) it is evident from the completeness of the Fourier integral that

$$\begin{aligned}
\mathcal{D}(x, y, z, \omega) &= \epsilon(x, y, z, \omega) \mathcal{E}(x, y, z, \omega), \\
\mathcal{J}(x, y, z, \omega) &= \sigma(x, y, z, \omega) \mathcal{E}(x, y, z, \omega), \\
\mathcal{H}(x, y, z, \omega) &= \mu^{-1}(x, y, z, \omega) \mathcal{B}(x, y, z, \omega),
\end{aligned}
\tag{27}$$

which is the usual connection of \mathcal{D} and \mathcal{J} to \mathcal{E} , and of \mathcal{H} to \mathcal{B} . Here one can see directly that $\mu^{-1}(x, y, z, \omega)$ is the reciprocal of the usual expression for magnetic permeability. It was in anticipation of this result that the notation $\hat{\mu}^{-1}(x, y, z, \tau)$ was used and that it was referred to earlier as representing instantaneous inverse magnetic permeability.

The input admittance $\mathcal{Y}(\omega)$ corresponding to the impulse response of current $\hat{y}(\tau)$ in (13) can now be obtained from the line integral of the connection (23) between $\hat{\mathcal{H}}$ and its

spectral equivalent \underline{H} ,

$$\oint \underline{H}(0,0,0,\omega) \cdot d\underline{l} = \int_0^\infty \left[\oint \hat{\underline{H}}(0,0,0,\tau) \cdot d\underline{l} \right] e^{-i\omega\tau} d\tau . \quad (28)$$

The line integral in brackets is $\hat{Y}(\tau)$, from (13), whereas the line integral on the left hand side represents the corresponding spectral function $Y(\omega)$. Hence

$$Y(\omega) = \int_0^\infty \hat{Y}(\tau) e^{-i\omega\tau} d\tau . \quad (29)$$

This familiar formula from circuit theory is usually stated as a definition for $Y(\omega)$. It has been derived here as a consequence of the two physical assumptions on which the present theory of systems is based.

5. SPECTRAL DESCRIPTION OF SIGNALS

It seems strange that the spectral description of electromagnetic systems should be widely recognized as possessing a physical origin and yet the spectral description of the signals associated with them be treated as though it had none. The meaning of signal spectrum has always been considered to be an arbitrary matter of definition rather than a necessary physical consequence of field theory. It is nearly always introduced into system theory by defining it to be the Fourier integral of the signal history over all time, including the future as well as the past,

$$V(\omega) = \int_{-\infty}^{\infty} v(t') e^{-i\omega t'} dt' . \quad (30)$$

Yet the particular course that the signal might take in the future can have no conceivable influence on anything that is happening now. Surely, then, any physically meaningful concept of signal spectrum must depend upon only the past history, changing from moment to moment as the record of the past

lengthens but always independent of what might happen in the future. Such considerations have led Page to suggest that it might be more appropriate to define an instantaneous spectrum (he called it a "running transform" [4(4)]) involving only the past. Again, however, he treated it as being purely a matter of definition. This presumed arbitrariness in the meaning of signal spectrum appears to have become so deeply ingrained as to lead A. A. Kharkevich, in an extensive treatment of the subject, to conclude the following [5]: "As we can see, the instantaneous spectrum can be defined in different ways. One thing should not be confused: all definitions are arbitrary. It is only necessary to choose the definition which is convenient for a given instance and formulate it clearly from the very beginning, in order to guarantee consistency throughout the discussion." It will be shown here that this is not entirely true. Quite the contrary, there is no arbitrariness at all. The instantaneous spectrum of any signal will be seen to be a definite, unique function that is completely determined physically at the point of observation by the instantaneous response of the system through which it is observed.

The instantaneous spectrum of an input signal comes directly from (7) or (9) after expressing the impulse response fields in terms of the corresponding spectral functions as in (22) and interchanging the order of integration. For the \underline{E} field, for example,

$$\begin{aligned}\underline{E}(x, y, z, t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\int_0^{t-\frac{z}{c}} \underline{v}(t') e^{-i\omega t'} dt' \right] \underline{E}(x, y, z, \omega) e^{i\omega t} d\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} V(\omega | t - \frac{z}{c}) \underline{E}(x, y, z, \omega) e^{i\omega t} d\omega ,\end{aligned}\tag{31}$$

where

$$V(\omega|t-\frac{r}{c}) \triangleq \int_0^{t-\frac{r}{c}} v(t') e^{-i\omega t'} dt' \quad (32)$$

The function $V(\omega|t-\frac{r}{c})$ is the full spectral description of the only piece of the input signal that produced the observable output ξ "now" at time t . That it depends on the input signal history up to time $t-\frac{r}{c}$ only, even though the signal history up to time t has already occurred, results from the fact that an observer at a distance r away from the source has absolutely no way of detecting any effect of the signal that occurred after $t-\frac{r}{c}$ because of the finite velocity c required for such effects to be communicated to him. It is evident, then, that the instantaneous spectrum $V(\omega|t-\frac{r}{c})$ represents the full instantaneous effect of the input signal upon the responses within a physical system as seen at a distance r away from the input.

The reason why the time-invariant function $V(\omega)$ in (30) is physically wrong is not that it doesn't represent the effect at (x, y, z, t) of the input signal, which it does, but because it contains as extra baggage the future signal as well:

$$V(\omega) = \int_{-\infty}^{t-\frac{r}{c}} v(t') e^{-i\omega t'} dt' + \int_{t-\frac{r}{c}}^{\infty} v(t') e^{-i\omega t'} dt'. \quad (33)$$

The second term in (33) involves only the future signal, which does not yet exist. To see why the signal spectrum can be disguised as being time-invariant by carrying along this extra term, look at the integral that represents the instantaneous output:

$$\begin{aligned} \frac{1}{2\pi} \int_{-\infty}^{\infty} V(\omega) \underline{\underline{E}}(x, y, z, \omega) e^{i\omega t} d\omega &= \int_{-\infty}^{t-\frac{\Delta}{c}} v(t') \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} \underline{\underline{E}}(x, y, z, \omega) e^{i\omega(t-t')} d\omega \right] dt' \\ &+ \int_{t-\frac{\Delta}{c}}^{\infty} v(t') \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} \underline{\underline{E}}(x, y, z, \omega) e^{i\omega(t-t')} d\omega \right] dt', \end{aligned} \quad (34)$$

after separating $V(\omega)$ into its past and future parts as in (31) and interchanging the order of integration. The bracketed inner integral represents the impulse response $\hat{\underline{\underline{E}}}(x, y, z, t-t')$. Hence the first term alone is the full representation for the $\underline{\underline{E}}$ field, the second term always being zero because of causality. The disguise was possible because the output is totally insensitive to the spectrum of the future input signal.

One might question whether the instantaneous spectrum (32) of an input signal must represent the instantaneous spectrum of any signal. Suppose one wanted the spectrum of the output rather than the input, what then? But the fact of the matter is that any signal must be the input to some physical system when it is being observed. Even if the sole interest were to observe its spectrum it still must be applied to the input of a spectrum analyzer. Hence it is concluded that the concept of signal spectrum has physical meaning only insofar as it represents the effect of an input signal on the instantaneous output of some physical system. Any purely abstract definition is devoid of physical reality.

It is of interest that the instantaneous spectrum (32) is essentially the same as Page's running transform but with the distinction that it appears here as a necessary consequence of the physical assumptions made earlier rather than as an intuitively appealing definition.

An important feature that distinguishes the instantaneous spectrum (32) from the time-invariant spectrum (30) is that the integral always converges. As a consequence there is no need for the convergence factor usually introduced for non-decaying signal functions such as the unit step function. This completely destroys one of the principal arguments favoring the Laplace over the Fourier transform in linear system theory.

6. EXACT RESPONSE TO A SINUSOIDAL INPUT SIGNAL

Now that the structure of the present theory of systems is complete the physical meaning of time-harmonic response on which much of system theory has been based in the past will be derived. Since no physical signal could have existed forever let the start of an input sinusoidal signal be the reference time $t' = 0$,

$$v(t') = \begin{cases} V_0 \cos(\omega_0 t' - \phi) & , \quad t' \geq 0 \\ 0 & , \quad t' < 0 \end{cases} \quad (35)$$

where V_0 and ω_0 are real positive numbers representing amplitude and angular frequency and ϕ is the starting phase angle. The responses at any time t can be obtained directly from (7) or (9) in terms of the impulse response fields in a form like that obtained for circuits by Carson [6] many years ago. For the \underline{E} field, for example,

$$\begin{aligned} \underline{E}(x, y, z, t) &= \int_0^{t-\frac{r}{c}} V_0 \cos(\omega_0 t' - \phi) \hat{\underline{E}}(x, y, z, t-t') dt' \\ &= V_0 \operatorname{Re} \left[\int_{-\frac{r}{c}}^t \hat{\underline{E}}(x, y, z, \tau) e^{i[\omega_0(t-\tau)-\phi]} d\tau \right] \\ &= V_0 \operatorname{Re} \left[\underline{E}(x, y, z, \omega_0) e^{i(\omega_0 t - \phi)} - e^{i(\omega_0 t - \phi)} \int_t^{\infty} \hat{\underline{E}}(x, y, z, \tau) e^{-i\omega_0 \tau} d\tau \right]. \end{aligned} \quad (36)$$

This is an exact expression for the \underline{E} field response everywhere in the system. At all times $t \leq \frac{r}{c}$ prior to arrival of the signal the second term exactly cancels the first, giving zero field as required by causality. This was illustrated in Fig. 5. The usual interpretation of the first term as steady-state response comes from the fact that the second term vanishes asymptotically with increasing t as a consequence of the exponentially decaying bound (14) on the impulse response, leaving only

$$\underline{E}(x, y, z, t) \underset{t \rightarrow \infty}{\sim} V_0 \operatorname{Re} [\underline{E}(x, y, z, \omega_0) e^{i(\omega_0 t - \phi)}] \quad (37)$$

This asymptotic steady-state behavior is the sole physical meaning that can be attributed to time-harmonic fields. It is limited to frequencies ω_0 that are purely real and positive, having no physical meaning at all for complex frequencies.

Some particularly interesting features of system theory appear when the response (36) is obtained from the spectral rather than the temporal descriptions of the system and the signal. The instantaneous spectrum of the input signal at the retarded time $t - \frac{r}{c}$, from (35) in (32), is

$$V(\omega | t - \frac{r}{c}) = \frac{i V_0}{2} \left[e^{-i\phi} \frac{e^{-i(\omega - \omega_0)(t - \frac{r}{c})} - 1}{\omega - \omega_0} + e^{i\phi} \frac{e^{-i(\omega + \omega_0)(t - \frac{r}{c})} - 1}{\omega + \omega_0} \right], \quad t \geq \frac{r}{c}. \quad (38)$$

It is a well behaved function, completely free of the singularities that occur at $\omega = \pm \omega_0$ in the artificially time-invariant spectrum $V(\omega)$. With (38) in (31) the instantaneous response of the system is found to be

$$\underline{E}(x, y, z, t) = V_0 \operatorname{Re} \left[\frac{e^{-i\phi}}{i 2\pi} \int_{-\infty}^{\infty} \frac{\underline{E}(x, y, z, \omega) e^{i\omega t}}{\omega - \omega_0} d\omega - \frac{e^{i[\omega_0(t - \frac{r}{c}) - \phi]}}{i 2\pi} \int_{-\infty}^{\infty} \frac{\underline{E}(x, y, z, \omega) e^{i\omega \frac{r}{c}}}{\omega - \omega_0} d\omega \right], \quad (39)$$

after recognizing from (23) that

$$\underline{E}(x, y, z, -\omega) = \underline{E}^*(x, y, z, \omega), \quad (40)$$

and hence that

$$\int_{-\infty}^{\infty} \frac{\underline{E}(x, y, z, \omega) e^{i\omega t}}{\omega + \omega_0} d\omega = - \int_{-\infty}^{\infty} \frac{\underline{E}^*(x, y, z, \omega) e^{-i\omega t}}{\omega - \omega_0} d\omega. \quad (41)$$

The result (39) is exactly equivalent to (36). It reduces to (36) when the two integrals are each expressed in terms of the impulse response, by using (23) and interchanging order of integration,

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{\underline{E}(x, y, z, \omega) e^{i\omega t}}{\omega - \omega_0} d\omega &= \int_{t/c}^{\infty} \hat{\underline{E}}(x, y, z, \tau) \left[\int_{-\infty}^{\infty} \frac{e^{i\omega(t-\tau)}}{\omega - \omega_0} d\omega \right] d\tau \\ &= i\pi \int_{t/c}^t \hat{\underline{E}}(x, y, z, \tau) e^{i\omega_0(t-\tau)} d\tau - i\pi \int_t^{\infty} \hat{\underline{E}}(x, y, z, \tau) e^{i\omega_0(t-\tau)} d\tau \\ &= i\pi e^{i\omega_0 t} \underline{E}(x, y, z, \omega_0) - i2\pi e^{i\omega_0 t} \int_t^{\infty} \hat{\underline{E}}(x, y, z, \tau) e^{-i\omega_0 \tau} d\tau. \quad (42) \end{aligned}$$

The bracketed integral on ω is the Hilbert transform of a complex exponential function,

$$\int_{-\infty}^{\infty} \frac{e^{i\omega(t-\tau)}}{\omega - \omega_0} d\omega = \begin{cases} i\pi e^{i\omega_0(t-\tau)} & , \quad \tau < t \quad (\text{past}) \\ 0 & , \quad \tau = t \quad (\text{present}) \\ -i\pi e^{i\omega_0(t-\tau)} & , \quad \tau > t \quad (\text{future}) \end{cases} \quad (43)$$

It is this abrupt change in the discontinuous integral (43) at the present time "now" that forced the τ -integration in (42) to break into past and future parts. Thus the Hilbert transform (43) is a key point in system theory. It is the fundamental link between the causal properties of the temporal and the spectral description of any physical system. As such it is the origin of the dispersion relations that appear throughout electromagnetic theory, resulting in the connection between the dispersive and absorptive properties of the medium ascribed to $\epsilon - \epsilon_0$, $\mu - \mu_0$, and σ and between the real and imaginary parts of \underline{E} , \underline{B} , \underline{D} , \underline{J} , \underline{H} for the system as a whole. For example, from (42),

$$\int_{-\infty}^{\infty} \frac{\underline{E}(x, y, z, \omega) e^{i\omega t}}{\omega - \omega_0} d\omega = -i\pi \underline{E}(x, y, z, \omega_0) e^{i\omega_0 t}, \quad t \leq \frac{1}{c} \quad (44)$$

which is the dispersion relation between the real and imaginary parts of $\underline{E}(x, y, z, \omega)$.

There seems to be a widespread misconception in the literature regarding the physical nature of the spectral functions \underline{E} , \underline{B} , \underline{D} , \underline{J} , and \underline{H} . They are frequently said to represent complex amplitudes of sinusoidally varying fields in space but with no reference to the system in which those fields must exist. The implication that they can represent arbitrary sinusoidally varying fields in space is contrary to the fact that they describe the fields of a specific system excited by a specific input signal (a unit impulse). It is important to recognize that their purpose is to describe systems, not fields. They are quite different from the temporal field functions \underline{E} , \underline{B} , \underline{A} , \underline{J} , and \underline{X} of Maxwell, which have no boundary conditions imposed upon them and for which even the medium itself need not have been specified.

7. CONCLUSION

Few, if any, of the results developed in this paper are really new. The importance of the development here is believed to lie not in the results themselves but in the demonstration that they are all derivable from experimentally verifiable assumptions. With these assumptions as a physical foundation the theory of electromagnetic systems can take its rightful place alongside other experimentally based physical theories.

REFERENCES

1. A. Einstein, "The Meaning of Relativity", Princeton University Press, Princeton, N. J., 1955; p.28.
2. R. P. Feynman, "The Character of Physical Law", M. I. T. Press, Cambridge, Mass., 1965; Chap. 5.
3. R. E. A. C. Paley and N. Wiener, "Fourier Transforms in the Complex Domain", American Math. Soc. Colloq. Pub., vol 19, 1934; p.2.
4. C. H. Page, "Instantaneous Power Spectra", Jour. App. Phys., vol 23, 1952, pp.103-106.
5. A. A. Kharkevich, "Spectra and Analysis", Consultants Bureau, N. Y., 1960; p.24.
6. J. R. Carson, "Theory of the transient oscillations of electrical networks and transmission systems", J. Am. Inst. Elect. Engr., vol 38, pp.345-427, 1919; eqns (13) and (14).

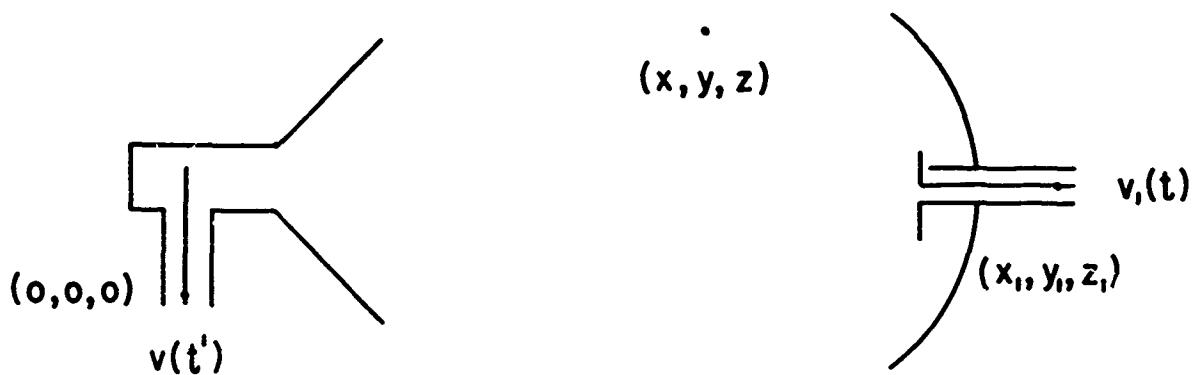


Fig. 1 - A two-port electromagnetic system consisting of two antennas. A voltage $v(t')$ at the input port $(0,0,0)$ produces a voltage $v_1(t)$ at the output port (x_1, y_1, z_1) .

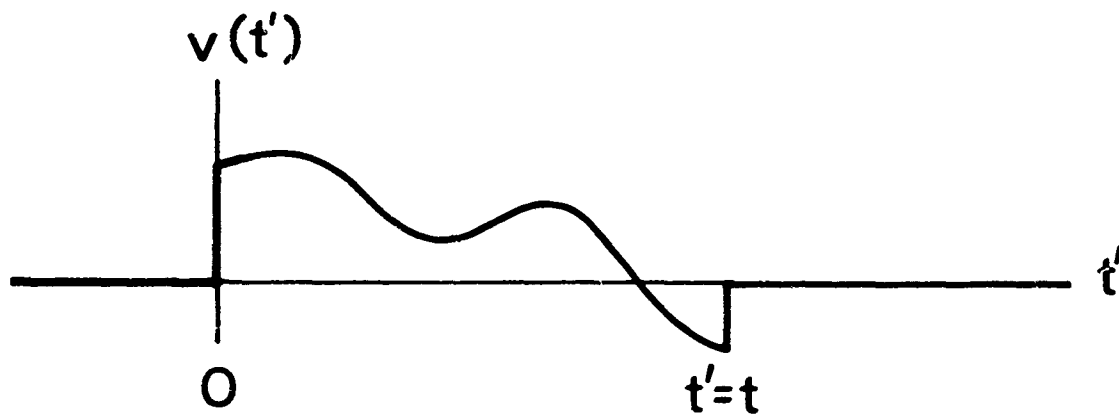


Fig. 2 - Physical development of a signal record as it exists at time t . Each infinitesimal element of time adds another pulse to the record.

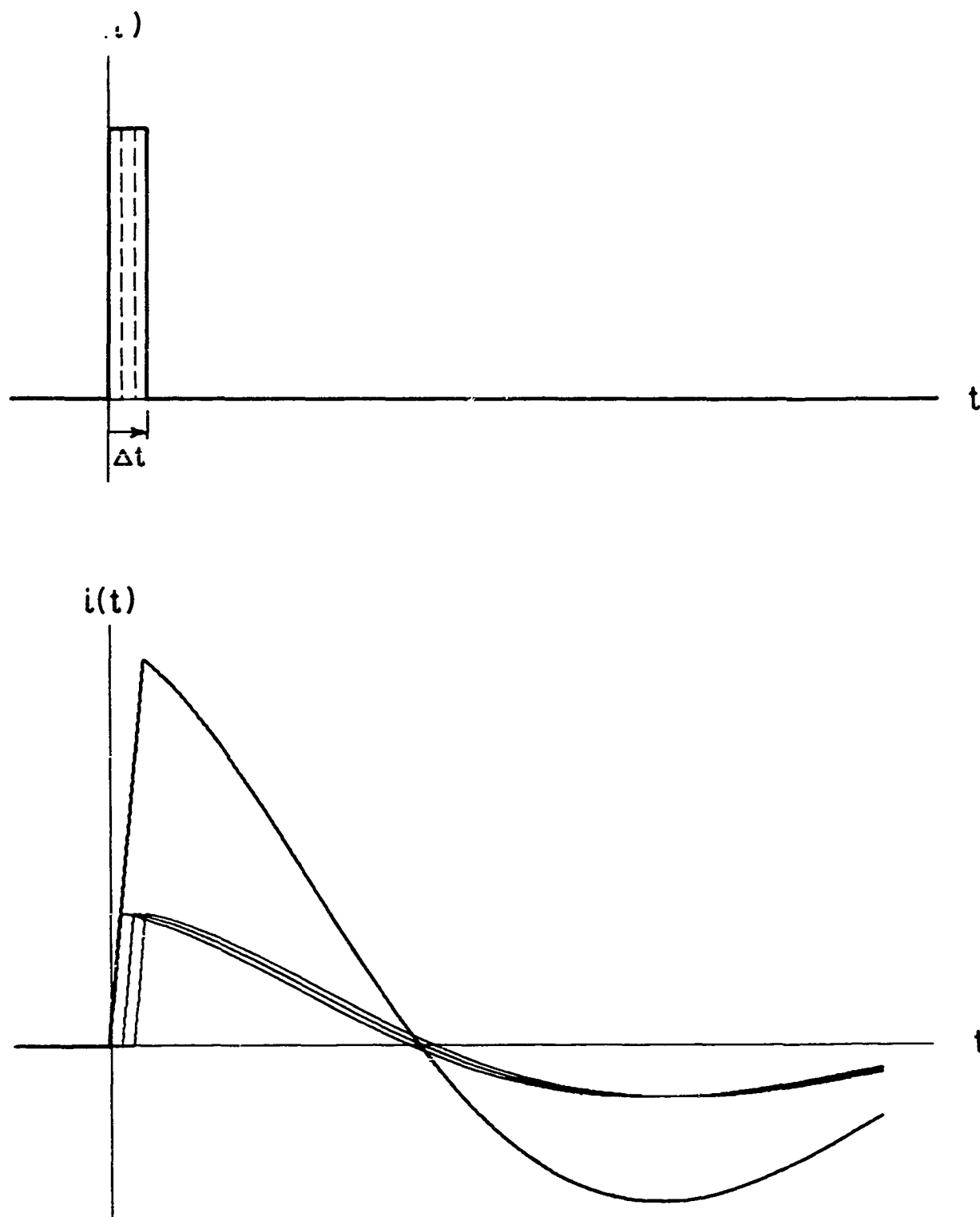


Fig. 3 - Input current response $i(t)$ to a rectangular voltage pulse $v(t')$ of duration Δt , and the individual responses to each of its n ($=3$) identical slices of duration $\Delta t/n$.

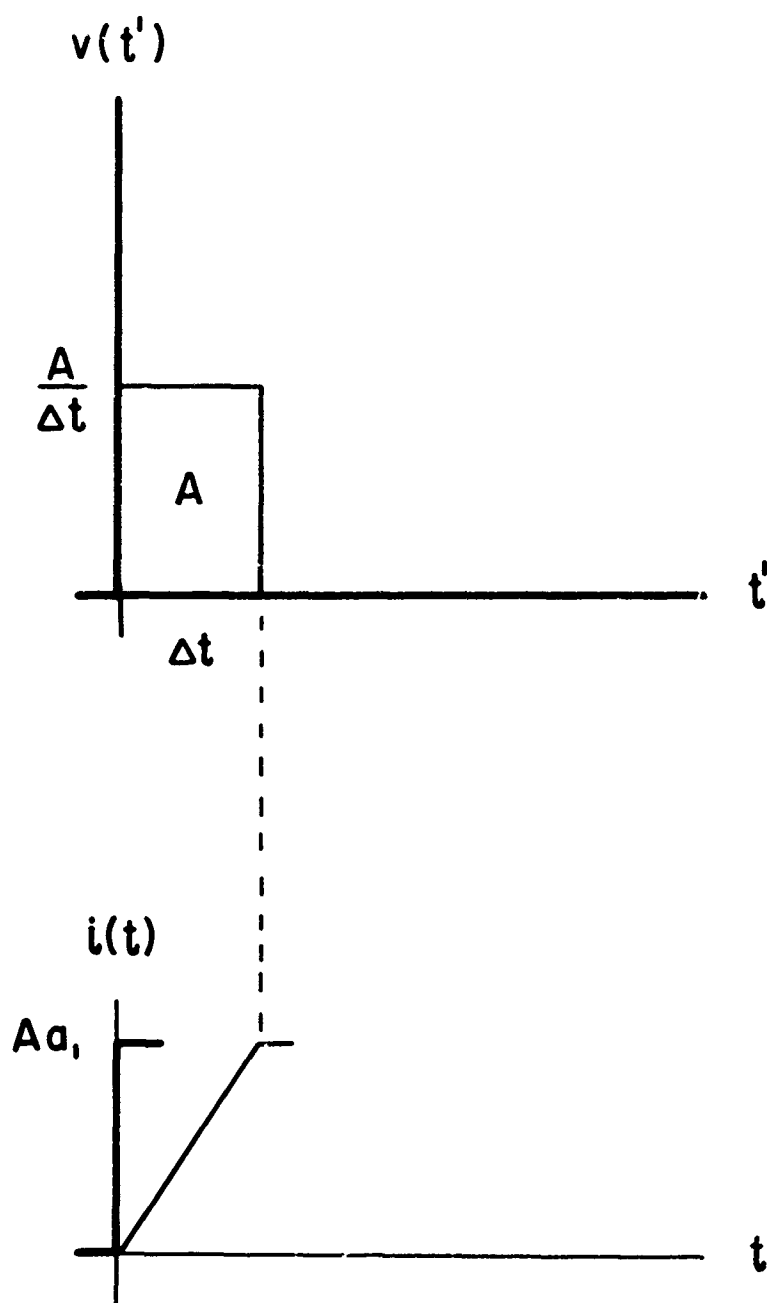


Fig. 4 - Initial input current response $i(t)$ to a rectangular voltage pulse of area A and vanishing duration Δt .

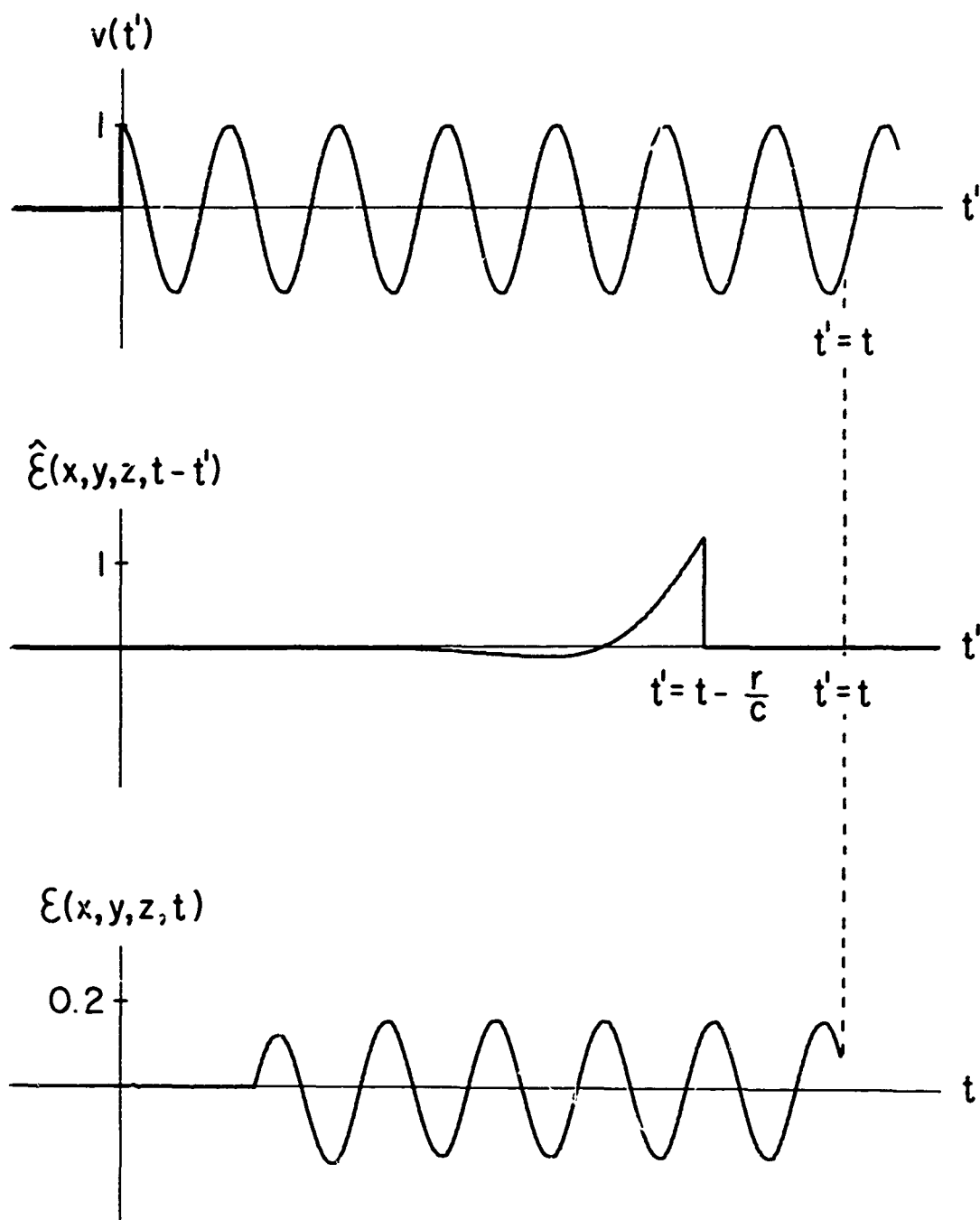


Fig. 5 - Relationship of the $\underline{\xi}$ -field in a system to the input voltage and a hypothetical $\hat{\underline{\xi}}$ -field response to a unit impulse of voltage. a) Sinusoidal voltage $v(t')$ at the input port $(0,0,0,)$ at time t' ; b) $\hat{\underline{\xi}}$ -field response at (x,y,z) at time t to a unit impulse of voltage at the input port at time t' ; c) resulting $\underline{\xi}$ -field produced at (x,y,z) at time t by the input voltage $v(t')$ as calculated by (7).

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) North Carolina State University Raleigh, North Carolina 27607		2a. REPORT SECURITY CLASSIFICATION Unclassified	
		2b. GROUP	
3. REPORT TITLE ON A PHYSICAL FOUNDATION FOR ELECTROMAGNETIC SYSTEM THEORY			
4. DESCRIPTIVE NOTES (Type of report and, inclusive dates) Scientific. Interim.			
5. AUTHOR(S) (First name, middle initial, last name) Donald R. Rhodes			
6. REPORT DATE 20 June 1969		7a. TOTAL NO. OF PAGES 38	7b. NO. OF REFS 6
8a. CONTRACT OR GRANT NO. F 19628-67-C-0255		8b. ORIGINATOR'S REPORT NUMBER(S) Scientific Report No. 1	
b. PROJECT, Task, Work Unit Nos. 4600-10-01			
c. DoD Element 62702F		9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) AFCRL - 69 - 0273	
d. DoD Subelement 674600			
10. DISTRIBUTION STATEMENT Distribution of this document is unlimited. It may be released to the Clearinghouse, Department of Commerce, for sale to the general public.			
11. SUPPLEMENTARY NOTES TECH, OTHER		12. SPONSORING MILITARY ACTIVITY Air Force Cambridge Research Laboratories (CRD) L. G. Hanscom Field Bedford, Massachusetts 01730	
13. ABSTRACT The traditional basis for the theory of electromagnetic systems (e.g., circuits, microwave networks, antennas) is a collection of definitions describing their temporal and spectral properties. It is shown here that these properties are all derivable as a consequence of two physical assumptions, namely, the Maxwell-Lorentz theory of electromagnetism and the postulates of causality, linearity, time-invariance, and passivity. These two assumptions alone thus constitute an axiomatic physical foundation for electromagnetic system theory.			

Unclassified

Security Classification

Unclassified
Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
System theory Signal Theory						

DD FORM 1 NOV 65 1473 (BACK)
S/N 0101-507-6821

Unclassified
Security Classification

A-31409