JOURNAL OF MATHEMATICAL PSYCHOLOGY 23, 1–89 (1981)

Lectures on Linear Systems Theory

M. FRANK NORMAN

Department of Psychology, University of Pennsylvania, 3813-15 Walnut Street, Philadelphia, Pennsylvania 19104

CONTENTS

PREFACE

In 1977 and 1979 the Mathematical Psychology Proseminar at the University of Pennsylvania was a one-semester course in Linear Systems Theory. I lectured on the general theory of linear systems, and several colleagues presented applications to selected psychological and psychophysiological problems. This article is a revision of my lecture notes. I am extremely grateful to E. S. Krendel, C. R. Gallistel, J. Nachmias, E. N. Pugh, and B. S. Rosner for their lectures and other contributions to the course, and to my thinking on this topic.

Both graduate and undergraduate students enrolled in the course. All had had at least a year of college calculus. No additional mathematical preparation was assumed, though I am certain that students with additional preparation (e.g., a third semester of calculus, or courses in linear algebra or physics) found it useful.

0022-2496/81/010001–89$02.00/0

In these lectures, certain statements are called "theorems," and the arguments supporting these statements are called "proofs." These terms are not meant to indicate that the exposition is mathematically rigorous. It is occasionally rigorous, but not consistently so.

Since few psychological applications of linear systems theory are discussed in this article, the reader will surely want to look elsewhere for illustrations of how these ideas are used in psychological models. The books of Harris (1980), McFarland (1971), Milsum (1966), and Stark (1968) are of interest in this connection, as are the *Journal of the Acoustical Society of America* and *Vision Research.* Let me say at the onset that I doubt that purely linear models will take us very far in the analysis of psychological phenomena. However, useful models may have linear components (hidden, perhaps, behind thresholds) or linear approximations (for, say, restricted ranges of inputs). Moreover, there is another reason for including "linear systems theory" in the psychological curriculum: the ideas grouped under this heading are so pervasive as to represent an important component of the language and culture of science.

PART A. ANALYSIS IN THE TIME DOMAIN

As a general reference on time domain analysis, I recommend Chapters 1 and 3 of Schwarz and Friedland (1965). Chapters 21–25 of Feynman, Leighton, and Sands (1963) show how linear systems ideas arise in elementary classical physics.

Lecture 1. *RC Circuits*

Let's begin with an example. Consider an electrical circuit containing a resistance, a capacitance, and a voltage source. Such a circuit is diagrammed in Fig. 1. R and C are constants, but the impressed voltage, $x(t)$, may vary with time, and this causes temporal variation in the voltage, $y(t)$, across the capacitor.

Our objective is to obtain a formula that expresses the "output," y, in terms of the "input," x. The standard method of doing this involves setting up a differential equation involving x and y. We begin with the identity

$$v_{13} = v_{12} + v_{23},$$

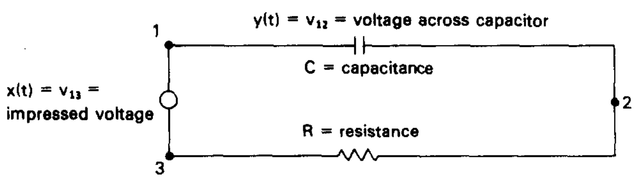

FIG. 1. An *RC* circuit.

which is a consequence of Kirchoff's second law. Replacing v_{13} and v_{12} by $x(t)$ and $y(t)$, respectively, we obtain

$$x(t) = y(t) + v_{23}.$$

Now v_{23} is the voltage across the resistance. According to Ohm's law,

$$v_{23} = i(t)R,$$

where $i(t)$ is the current in the circuit at time t. Thus

$$x(t) = y(t) + i(t)R. \tag{1}$$

I will now show that i is proportional to $y' = dy/dt$. Let

$$q(t) = \text{charge on capacitor at time } t.$$

Then

$$y(t) = q(t)/C.$$

(Think of y as the water pressure at the bottom of a vertical pipe containing q gallons of water. Larger pipes yield smaller pressure.) Thus

$$y'(t) = q'(t)/C$$

or

$$q'(t) = Cy'(t).$$

But

$$q'(t) = i(t),$$

so

$$i(t) = Cy'(t). \tag{2}$$

Thus i is proportional to y', as claimed.

Substitution of (2) into (1) yields

$$x(t) = y(t) + RCy'(t).$$

Interchanging the two sides of this equation, suppressing t, and introducing the new parameter

$$b = RC,$$

called the *time constant* of the circuit, we obtain

$$by' + y = x. \tag{3}$$

This is the differential equation that we must solve in order to express output, y, in terms of input, x.

Solution of (3)

We begin by observing that we know a solution of (3) in one special case, namely, $x(t) \equiv 0$. In that case, (3) reduces to $by' + y = 0$, and

$$p(t) = e^{-t/b}$$

is a solution. In other words,

$$bp' + p = 0.$$

The fact that this equation is so similar to (3) suggests that the unknown solution, y, of (3) may bear some simple relationship to p. Perhaps the ratio, y/p, is a simple function, or, at any rate, satisfies a very simple differential equation. Using the "quotient rule" to calculate its derivative, we obtain

$$\begin{aligned}(y/p)' &= \frac{py' - yp'}{p^2} \\ &= \frac{p(-b^{-1}y + b^{-1}x) - y(-b^{-1}p)}{p^2} \\ &= \frac{-b^{-1}py + b^{-1}px + b^{-1}yp}{p^2}.\end{aligned}$$

Cancelling $b^{-1}py$ in the numerator, we get

$$(y/p)' = b^{-1}px/p^2$$

or

$$(y/p)' = b^{-1}x/p.$$

This differential equation is very simple indeed, since the right hand side is a known function of t. Integrating both sides from t_0 to t, we obtain

$$\int_{t_0}^{t} (y(u)/p(u))'\, du = b^{-1} \int_{t_0}^{t} (x(u)/p(u))\, du.$$

Hence, by the fundamental theorem of calculus,

$$y(t)/p(t) - y(t_0)/p(t_0) = b^{-1} \int_{t_0}^{t} (x(u)/p(u))\, du.$$

Therefore

$$y(t) = p(t)\left[y(t_0)/p(t_0) + b^{-1}\int_{t_0}^{t} (x(u)/p(u))\,du\right]. \tag{4}$$

This formula expresses y in terms of the input function, x, the known function $p(t) = e^{-t/b}$, and the *initial value* $y(t_0)$. The differential equation does not determine the initial value. We can choose it as we please to describe different initial voltages across the capacitor. However, once $y(t_0)$ is specified, all other values of $y(t)$ are determined and are given by the above formula.

The voltage inputs, $x(t)$, that we might actually apply to an RC circuit have a definite *onset time*, t_x, illustrated in Fig. 2. Clearly t_x can be described as the largest time with the property that the input is "off" at all earlier times ($x(t) = 0$ for all $t < t_x$). One can also imagine a *null input* that never comes on ($x(t) = 0$ for all t, hence $t_x = \infty$), and inputs, like $x(t) = \sin t$, which have been on since $t_x = -\infty$.

For inputs with finite onset times, we assume that there is no voltage across the capacitor before input voltage onset or, equivalently, $t_y \geqslant t_x$. This condition yields the initial value $y(t_0) = 0$ for all $t_0 < t_x$, so (4) reduces to

$$y(t) = b^{-1}p(t)\int_{t_0}^{t} (x(u)p(u))\,du$$

for $t_0 < t_x$. But $x(u) = 0$ for $u < t_0$, so we change nothing on the right if we extend the limit of integration from t_0 down to $-\infty$. Thus

$$y(t) = b^{-1}p(t)\int_{-\infty}^{t} (x(u)/p(u))\,du$$

or

$$y(t) = b^{-1}\int_{-\infty}^{t} (p(t)/p(u))\,x(u)\,du.$$

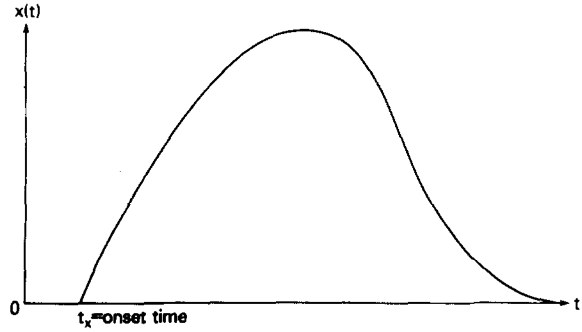

FIG. 2. Onset time, t_x.

But

$$\frac{p(t)}{p(u)} = \frac{e^{-t/b}}{e^{-u/b}} = e^{-(t-u)/b},$$

so

$$y(t) = b^{-1} \int_{-\infty}^{t} e^{-(t-u)/b} x(u)\, du. \tag{5}$$

This formula gives the voltage "output" of an initially uncharged capacitor for an input voltage with finite onset time.

EXAMPLE: UNIT STEP INPUT. Suppose that

$$\begin{aligned} x(t) &= 1, \qquad \text{if} \quad t \geqslant 0, \\ &= 0, \qquad \text{if} \quad t < 0. \end{aligned}$$

This is called the *unit step* input, and the corresponding output, y, is the *step response*. According to (5), $y(t) = 0$ for $t < 0$. For $t \geqslant 0$,

$$\begin{aligned} y(t) &= b^{-1} e^{-t/b} \int_{-\infty}^{t} e^{u/b} x(u)\, du \\ &= b^{-1} e^{-t/b} \int_{0}^{t} e^{u/b}\, du \\ &= b^{-1} e^{-t/b} \left. \frac{e^{u/b}}{1/b} \right|_{0}^{t} \\ &= e^{-t/b}(e^{t/b} - 1), \end{aligned}$$

so

$$y(t) = 1 - e^{-t/b}$$

for $t \geqslant 0$. This function is graphed in Fig. 3. Eventually the entire impressed voltage appears across the capacitor.

Infinite Onset Times

For theoretical purposes, it is also necessary to consider inputs with infinite onset times. The null input ($x(t) = 0$ for all t, $t_x = \infty$) presents no difficulty. We have *assumed* that there is no voltage across the capacitor before the input voltage is applied. In the case at hand, this implies $y(t) = 0$ for all t, which also follows from (5) if we take $x(u) = 0$ on the right. Thus (5) is valid for $t_x = \infty$ as well as for finite t_x.

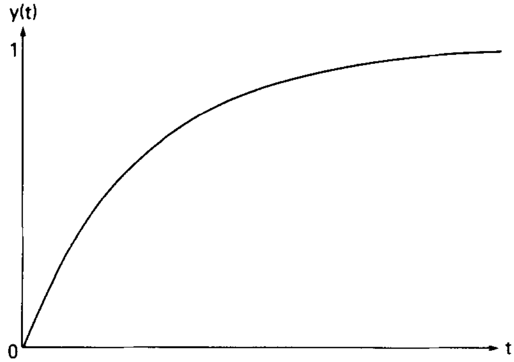

FIG. 3. Step response of an RC circuit.

Inputs that have been on since $t_x = -\infty$ require more careful consideration. It is easy to concoct an input of this kind for which the integral on the right in (5) is infinite, as you will see in the third exercise on p. 8. Fortunately, the integral is finite for a variety of functions, like $x(u) = \sin \omega u$, that are of special interest to us. It can be shown that, whenever the integral is finite, the function, $y(t)$, defined by (5) is a solution of the differential equation, (3), of the RC circuit. It is, in fact, the unique solution satisfying an appropriate "initial condition," the precise form of which need not concern us. (For the record, the condition I have in mind is $e^{t/b}y(t) \to 0$ as $t \to -\infty$.) Thus, even for certain "theoretical" inputs with $t_x = -\infty$, we can regard the function $y(t)$ of (5) as the corresponding "ideal" output of the RC circuit.

In summary, (5) gives the input–output relation for the RC circuit, regardless of the value of onset time, t_x.

EXAMPLE: SINUSOIDAL INPUT. We shall see that it is very important to know how a system of this kind responds to sinusoidal inputs. Thus, consider $x(t) = \sin \omega t$, where ω is frequency in radians/unit time. This input has $t_x = -\infty$. According to (5), the output is

$$\begin{aligned} y(t) &= b^{-1}e^{-t/b}\int_{-\infty}^{t} e^{u/b} \sin \omega u \, du \\ &= b^{-1}e^{-t/b} \lim_{T \to -\infty} \int_{T}^{t} e^{u/b} \sin \omega u \, du. \end{aligned} \tag{6}$$

We shall always interpret integrals with infinite endpoints as limits of integrals with finite endpoints. Consulting a table of indefinite integrals, we find

$$\int e^{ax} \sin bx \, dx = e^{ax} \frac{(a \sin(bx) - b \cos(bx))}{a^2 + b^2}.$$

Taking "a" $= b^{-1}$ and "b" $= \omega$, and subtracting the value of this expression at $x = T$ from its value at $x = t$, we obtain

$$\int_T^t e^{u/b} \sin \omega u \, du = e^{t/b} \frac{(b^{-1} \sin \omega t - \omega \cos \omega t)}{b^{-2} + \omega^2} - e^{T/b} \frac{(b^{-1} \sin \omega T - \omega \cos \omega T)}{b^{-2} + \omega^2}.$$

Now $e^{T/b} \to 0$ as $T \to -\infty$, and $\sin \omega T$ and $\cos \omega T$ oscillate between + and -1 as $T \to -\infty$. Thus the entire second term on the right converges to 0 as $T \to -\infty$. Therefore

$$\lim_{T \to -\infty} \int_T^t e^{u/b} \sin \omega u \, du = e^{t/b} \frac{(b^{-1} \sin \omega t - \omega \cos \omega t)}{b^{-2} + \omega^2}.$$

Plugging this into (6) and cancelling $e^{t/b}$, we obtain

$$y(t) = \frac{b^{-1}}{b^{-2} + \omega^2} (b^{-1} \sin \omega t - \omega \cos \omega t).$$

As we shall see later, the right hand side represents a sinusoid with frequency ω. Thus a sinusoidal input produces a sinusoidal output of the same frequency. Input and output differ only in amplitude and phase.

Exercises

Find t_x and $y(t)$ for inputs (a), (b), and (c).

(a)
$$x(u) = 1, \quad \text{if} \quad 0 \leqslant u \leqslant c,$$
$$= 0, \quad \text{if} \quad u < 0 \text{ or } u > c.$$

When calculating $y(t)$, consider separately the cases $t < 0$, $0 \leqslant t \leqslant c$, and $t > c$.

(b)
$$x(u) = \cos \omega u.$$

(c)
$$x(u) = e^{ku}.$$

Consider separately the cases $k > -b^{-1}$, $k = -b^{-1}$, and $k < -b^{-1}$.

Lecture 2. From RC Circuits to Linear Systems

In Lecture 1 we obtained the following basic formula for the voltage, $y(t)$, across a capacitor in an RC circuit as a function of the input voltage, $x(t)$:

$$y(t) = b^{-1} \int_{-\infty}^t e^{-(t-u)/b} x(u) \, du.$$

I prefer not to have to write

$$b^{-1}e^{-(t-u)/b}$$

again and again, so I am going to introduce a simpler notation for this quantity. Let

$$g(t) = b^{-1}e^{-t/b}. \tag{1}$$

Then

$$g(t-u) = b^{-1}e^{-(t-u)/b},$$

so that

$$y(t) = \int_{-\infty}^{t} g(t-u)\,x(u)\,du. \tag{2}$$

Equation (2) shows how the input, x, is transformed into an output, y. Note that the output at time t depends on all prior values of the input, not just on the input at time t. Most interesting physical, physiological, and psychological systems have this property. Thus (2) defines a mapping from *functions* (inputs) *to functions* (outputs), rather than from *numbers to numbers*. In other words, (2) defines a *function*, G, that maps functions (x) into functions (y). The calculus-style notation for this state of affairs would be something like $y = G(x)$, where, when I write "y" instead of "$y(t)$," I refer to the entire voltage function, not just to its value at time t. (In common usage, the notation "$y(t)$" is ambiguous, sometimes standing for the function, y, and sometimes for its value at t.)

Now the notion of "functions of functions" is potentially confusing, at least when considered in the abstract, so, instead of calling G a function, we call it a *transformation*, or an *operator*, or a *system*. Moreover, it is customary to write "Gx" instead of "$G(x)$." $y = Gx$ is a function of time, and we denote its value at time t in the usual way, $y(t) = Gx(t)$.

In summary, the transformation G is defined by $Gx = y$ or

$$Gx(t) = \int_{-\infty}^{t} g(t-u)\,x(u)\,du, \tag{3}$$

where

$$g(t) = b^{-1}e^{-t/b}. \tag{1}$$

Our next order of business is to show that the transformation G has many nice properties. All of these properties could be described directly in terms of y's and x's, without mention of G. However, the G notation helps to focus attention on fundamental relationships, and permits us to state these relationships more compactly, which, in turn, makes them easier to learn and remember.

Linearity

The first property of G is *linearity*. To understand what this is all about, we must understand how to add two functions, and how to multiply a function by a constant factor. This is very straightforward, but it is so basic that it is worth dignifying by a formal definition.

DEFINITION. Let x and y be functions (of time) and let a be a constant. Then the sum, or superposition, of x_1 and x_2 is the function, $x_1 + x_2$, given by

$$(x_1 + x_2)(t) = x_1(t) + x_2(t) \qquad \text{(for all } t\text{)}.$$

And the product, ax, of a and x is given by

$$(ax)(t) = a \times x(t).$$

Functions are added "pointwise." You can think of a in ax as an amplification factor. The graph of $2x$ looks like the graph of x, but all the vertical dimensions are doubled.

I claimed above that G was a *linear transformation* or *linear system*. This means that it has the following properties. First, the output produced by a superposition of inputs is just the superposition of the corresponding outputs. Symbolically,

$$G(x_1 + x_2) = Gx_1 + Gx_2. \tag{L1}$$

This is called the *superposition property*. The second property of linear systems is that amplification of the input produces the same amplification of the output,

$$G(ax) = aGx. \tag{L2}$$

This is called *homogeneity*.

The following proof may deepen your understanding of superposition.

Proof of (L1). The value of $G(x_1 + x_2)$ at time t is

$$[G(x_1 + x_2)](t) = \int_{-\infty}^{t} g(t-u)[(x_1 + x_2)(u)]\, du,$$

by the definition, (3), of G;

$$= \int_{-\infty}^{t} g(t-u)[x_1(u) + x_2(u)]\, du,$$

by the definition of $x_1 + x_2$;

$$= \int_{-\infty}^{t} [g(t-u)\, x_1(u) + g(t-u)\, x_2(u)]\, du,$$

by the distributive law relating multiplication and addition of real numbers ($a(b+c)=ab+ac$);

$$=\int_{-\infty}^{t} g(t-u)\,x_1(u)\,du+\int_{-\infty}^{t} g(t-u)\,x_2(u)\,du,$$

by a property of integration ($\int(f_1+f_2)=\int f_1+\int f_2$);

$$=Gx_1(t)+Gx_2(t),$$

by the definition of G;

$$=(Gx_1+Gx_2)(t),$$

by the definition of "+" for functions. Thus we have shown that

$$G(x_1+x_2)(t)=(Gx_1+Gx_2)(t).$$

In other words the functions (outputs) $G(x_1+x_2)$ and Gx_1+Gx_2 agree at every point in time. This, however, is what (L1) means!

The proof that G has property (L2) is similar, so I leave it to you as an exercise.

Invariance

The system G has another property that is every bit as important as linearity. This property is *invariance* or *stationarity*. To formulate this property economically, we must first discuss what it means to shift a function in time. For any function, x, and any number, τ (tau), let $T_\tau x$ be the function whose value at t is the value of x at $t-\tau$,

$$T_\tau x(t)=x(t-\tau). \tag{4}$$

If τ is positive, then T_τ introduces a delay of duration τ in the evolution of x. The graph of $T_\tau x$ is just the graph of x, shifted to the right [sic] by time τ, as in Fig. 4. Note that T_τ, like G, maps functions into functions. Thus it is referred to as the *shift transformation* or the *shift operator*.

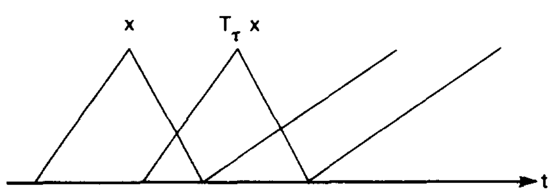

FIG. 4. Shift to the right.

To say that our system, G, is *invariant* means that the entire effect of shifting the input is to produce an equal shift in the output. Symbolically,

$$G(T_\tau x) = T_\tau(Gx). \tag{I}$$

On the left, we shift the input. On the right, we shift the output. Property (I) says we get the same function either way. Invariance does not mean that the output is constant. It means that the *mechanism* for input–output conversion is constant. In the RC circuit, this reduces to the stipulation that resistance, R, and capacitance, C, are constant over time.

Proof of (I). At time t, the left side is

$$[G(T_\tau x)](t) = \int_{-\infty}^{t} g(t-u)\, T_\tau x(u)\, du,$$

by the definition, (3), of G;

$$= \int_{-\infty}^{t} g(t-u)\, x(u-\tau)\, du,$$

by the definition, (4), of T_τ. We now introduce a new variable of integration, $w = u - \tau$. Solving for u in terms of w, we get $u = w + \tau$. Substituting into the integral, we obtain

$$\begin{aligned}\int_{-\infty}^{t} g(t-u)\, x(u-\tau)\, du &= \int_{-\infty}^{t-\tau} g(t-(w+\tau))\, x(w)\, \frac{du}{dw}\, dw \\ &= \int_{-\infty}^{t-\tau} g(t-\tau-w)\, x(w)\, dw \\ &= Gx(t-\tau) \\ &= [T_\tau(Gx)](t).\end{aligned}$$

Thus

$$[G(T_\tau x)](t) = [T_\tau(Gx)](t)$$

for all t, which is what property (I) means.

Nonanticipation

The system G has one additional property that is worth noting at this time. The output at time t,

$$Gx(t) = \int_{-\infty}^{t} g(t-u)\, x(u)\, du,$$

depends only on the input *up to* time t. Hence, if two inputs agree at least until time T, then their outputs also agree until time T. The system does not anticipate future inputs. It is said to be *nonanticipating* or *causal*,

$$x(t) = y(t) \text{ for } t < T \text{ implies } Gx(t) = Gy(t) \text{ for } t < T. \tag{N}$$

This property is, of course, quite natural and intuitive. It turns out that it is a direct consequence of the initial condition that we imposed on the solution to our differential equation. Indeed, the property is actually equivalent to the initial condition, for linear systems. Recall that our initial condition was that there be no voltage across the capacitor before the input voltage onset time, t_x. This can be stated as

$$Gx(t) = 0 \text{ for } t < t_x, \text{ or, letting } y = Gx,\ t_y \geqslant t_x. \tag{N$'$}$$

The similarity of (N) and (N′) is obvious, and, in fact, *(N) and (N′) are equivalent for any linear system G.* I will give the simple proof in Lecture 6.

Summary and Generalization

This is a good place to take stock of what we have done. We have shown that the input–output transformation, G, for the RC circuit is linear, invariant, and nonanticipating. All of our arguments have proceeded from the equation

$$Gx(t) = \int_{-\infty}^{t} g(t-u)\, x(u)\, du. \tag{3}$$

For the RC circuit, g has the form

$$g(t) = b^{-1} e^{-t/b}. \tag{1}$$

This may have slipped your mind, since we have not made any use of the equation for g in today's lecture. Thus we have really proved the following theorem, which goes far beyond RC circuits.

THEOREM. *For any function g, the system G defined by* (3) *is linear, invariant, and nonanticipating.*

In fact, this result holds not only for bona fide functions, g, but for certain generalized functions, which we will introduce next time. We shall also prove the converse. Essentially any linear, invariant, nonanticipating system G can be represented in the form (3) for a suitable function or generalized function g. The universality of (3) is the fundamental theorem of time-domain analysis of linear systems.

Exercises

1. Prove that the system defined by (3) satisfies (L2).
2. Consider the systems, G, defined by the following equations.
   (a) $Gx = 8x$,
   (b) $Gx(t) = tx(t)$,
   (c) $Gx(t) = x(t) + 15$,
   (d) $Gx(t) = (x(t))^2$,
   (e) $Gx = dx/dt$,
   (f) $Gx(t) = \int_0^t x(u)\, du$,
   (g) $G = T_u =$ shift by u, $u \geqslant 0$,
   (h) $G = T_u$, $u < 0$.

For each system, and each of the properties (L1), (L2), (I), and (N), prove that the system has the property or show by example that it does not. Here is one more, for good measure:

(i) $Gx(t) = x(2t)$.

Lecture 3. *Impulse and Impulse Response*

Last time we showed that a system, G, given by an equation of the form

$$Gx(t) = \int_{-\infty}^{t} g(t-u)\, x(u)\, du \tag{1}$$

is linear, invariant, and nonanticipating. Our main objective today is to prove the converse: any linear, invariant, nonanticipating system G can be represented in the form (1), for a uniquely determined function g.

The first step in this direction is to show how g can be recovered from G. Let me restate this problem in an amplified form: Suppose that G corresponds to a machine, hidden in one of the famous black boxes. We can perform a behavioral analysis of the machine. That is, we can stimulate it with any input and observe the resulting output, but we can't open the box. Suppose that we hypothesize that there is a function g such that input and output are related by (1). How should we stimulate the machine in order to determine g? The answer may surprise you. *If we kick box, the output will be g*!

Naturally, this statement requires interpretation. The "kick" I have in mind is a very short, very intense pulse, with

$$\text{strength} \times \text{duration} = 1. \tag{2}$$

The mathematical description of such a kick is rather amusing. We begin with an input pulse, x_ϵ, of duration ε and strength $1/\varepsilon$, so that (2) is satisfied,

$$\begin{aligned} x_\epsilon(t) &= 1/\varepsilon, \qquad \text{if} \quad 0 \leqslant t \leqslant \varepsilon, \\ &= 0, \qquad \text{otherwise.} \end{aligned}$$

This pulse is illustrated in Fig. 5. Here ε is an arbitrary positive number, but you can think of it as small. As $\varepsilon \to 0$, x_ϵ approaches the "kick."

What is the response of G to the input x_ϵ? To answer this, we begin by noting that, for any input x, Gx can be rewritten

$$Gx(t) = \int_{-\infty}^{\infty} g(t-u)\, x(u)\, du, \tag{1'}$$

provided that $g(\tau)$ is defined to be zero on the negative axis,

$$g(\tau) = 0 \qquad \text{for} \quad \tau < 0. \tag{3}$$

For if $u > t$ in the integral, then $g(t-u) = 0$ and (1′) reduces to (1). Granting this,

$$\begin{aligned} Gx_\epsilon(t) &= \int_{-\infty}^{\infty} g(t-u)\, x_\epsilon(u)\, du \\ &= \int_0^\epsilon g(t-u)\, x_\epsilon(u)\, du, \end{aligned}$$

since $x_\epsilon(u) = 0$ except for $0 \leqslant u \leqslant \varepsilon$, hence

$$Gx_\epsilon(t) = \varepsilon^{-1} \int_0^\epsilon g(t-u)\, du. \tag{4}$$

So far we have assumed nothing about g (except (3)). Let us now assume that g is continuous at t. This means that $g(\tau) \to g(t)$ as $\tau \to t$, or that the graph of g does not have a break at t. In Fig. 6, g is continuous at all points, t, except $t = 0$ and $t = 13.3$.

Returning to (4), we see that, when ε is small, $t-u$ is close to t, so $g(t-u)$ is close to $g(t)$, by continuity. Thus, as $\varepsilon \to 0$, we can replace $g(t-u)$ in (4) by $g(t)$, which yields

$$\lim_{\epsilon \to 0} Gx_\epsilon(t) = \lim_{\epsilon \to 0} \varepsilon^{-1} \int_0^\epsilon g(t)\, du = g(t)\, \varepsilon_{-1} \int_0^\epsilon du = g(t).$$

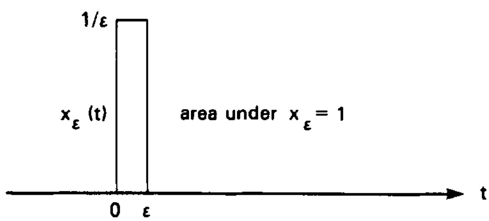

FIG. 5. The input pulse x_ϵ.

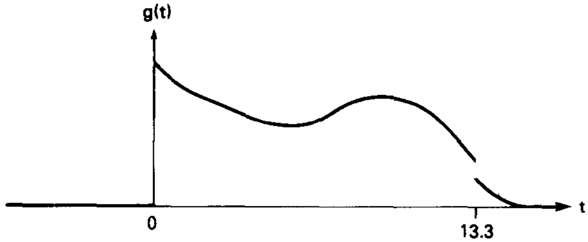

FIG. 6. A discontinuous function.

Thus

$$\lim_{\epsilon \to 0} Gx_\epsilon(t) = g(t), \tag{5}$$

provided only that g is continuous at t.

Now I hope you will agree that this is straightforward and in no way mysterious. However, the same cannot be said for the notation we will now introduce,

$$\delta = \text{unit impulse} = \text{delta function.}$$

This is just the "kick" that we have been discussing. It is supposed to be the limit of x_ϵ as $\varepsilon \to 0$,

$$\lim_{\epsilon \to 0} x_\epsilon = \delta.$$

The meaning of this limit is not too clear, but, interestingly enough, the limit in (5) is perfectly well defined for most interesting systems, and we can, in turn, use this limit to define $G\delta$, which we interpret as the response of the system to an instantaneous unit impulse.

DEFINITION. $G\delta(t) = \lim_{\epsilon \to 0} Gx_\epsilon(t)$. This quantity is called the *impulse response* of G.

We can then interpret the argument leading up to (5) as a proof of the following theorem.

THEOREM 1. *If a linear system G is given by*

$$Gx(t) = \int_{-\infty}^{\infty} g(t-u)\, x(u)\, du, \tag{1'}$$

then $G\delta = g$, i.e., g is the impulse response of G.

More about δ

The argument that led to (5) shows that, if a function h is continuous at 0, then

$$\lim_{\epsilon \to 0} \int_{-\infty}^{\infty} h(u)\, x_\epsilon(u)\, du = h(0). \tag{6}$$

Just as above, we might consider the left hand side to be a definition of $\int_{-\infty}^{\infty} h(u)\, \delta(u)\, du$. In view of (6), this is equivalent to the following definition of this integral

$$\int_{-\infty}^{\infty} h(u)\, \delta(u)\, du = h(0). \tag{7}$$

The conventional viewpoint is that this equation is not a definition of the whole expression on the left, but is, instead, a definition of δ. The problem with this approach is that it is easy to show that there isn't *any* bona fide function that satisfies (7)! Thus δ is called a ***generalized function.*** It is something of a miracle that the theory of such generalized functions really works, and can, in fact, be put on a completely rigorous mathematical footing via high-powered methods. In particular, most of the standard recipes found in calculus books work just as well for generalized functions as for ordinary functions. We shall make use of this fact later.

At any rate, once we are in possession of (7) we can obtain a marvelously direct proof of the last theorem. For if

$$Gx(t) = \int_{-\infty}^{\infty} g(t-u)\, x(u)\, du,$$

then, taking $x = \delta$, we get

$$\begin{aligned} G\delta(t) &= \int_{-\infty}^{\infty} g(t-u)\, \delta(u)\, du \\ &= g(t-0) = g(t). \end{aligned}$$

So $G\delta = g$, as claimed.

Digression on Superposition

Suppose that G is a system that has the superposition property,

$$G(x_1 + x_2) = Gx_1 + Gx_2. \tag{L1}$$

Nothing else is assumed about G. It follows from (L1) that

$$G(x_1 + x_2 + x_3) = Gx_1 + Gx_2 + Gx_3.$$

For

$$G(x_1 + x_2 + x_3) = G((x_1 + x_2) + x_3)$$
$$= G(x_1 + x_2) + Gx_3,$$

by (L1);

$$= Gx_1 + Gx_2 + Gx_3,$$

by a second application of (L1).

Similarly, $n-1$ applications of (L1) yield

$$G\left(\sum_{j=1}^{n} x_j\right) = \sum_{j=1}^{n} Gx_j \tag{L1}$$

for any $n \geqslant 1$. We recover the original version of (L1) by taking $n = 2$. It is but a short step from finite sums to infinite sums, and we will sometimes wish to assume that our system has the stronger superposition property

$$G\left(\sum_{j=1}^{\infty} x_j\right) = \sum_{j=1}^{\infty} Gx_j. \tag{L1$'$}$$

This arises, for example, when we attempt to apply G to an input, x, represented by a Fourier series

$$x(t) = \sum_{j=1}^{\infty} (a_j \cos jt + b_j \sin jt).$$

Here

$$x_j(t) = a_j \cos jt + b_j \sin jt.$$

We shall see that, if x is not periodic, we must superpose a continuum of exponentials to obtain a Fourier integral representation,

$$x(t) = \int_{-\infty}^{\infty} (a(\omega) \cos \omega t + b(\omega) \sin \omega t)\, d\omega$$

or

$$x = \int_{-\infty}^{\infty} x_\omega \, d\omega,$$

where

$$x_\omega(t) = a(\omega) \cos \omega t + b(\omega) \sin \omega t.$$

The superposition property for such "continuous sums" is

$$G\left(\int_{-\infty}^{\infty} x_\omega \, d\omega\right) = \int_{-\infty}^{\infty} Gx_\omega \, d\omega. \tag{L1$''$}$$

When the need arises, we shall assume that our "linear systems" possess this extended superposition property.

Main Theorem

You will recall that our main objective was to obtain the following result.

THEOREM 2. *A linear, invariant, nonanticipating system G admits the representation*

$$Gx(t) = \int_{-\infty}^{t} g(t-u)\,x(u)\,du, \tag{1}$$

where $g = G\delta$.

Proof. The unit impulse, δ, is a spike at time 0. It follows that $\delta(t-u)$, regarded as a function of u with t fixed, is a spike at time t ($t-u=0$ when $u=t$). Thus, by analogy with (7),

$$\int_{-\infty}^{\infty} x(u)\,\delta(t-u)\,du = x(t).$$

But $\delta(t-u)$, regarded as a function of t with u fixed, is just the δ function, shifted by u,

$$\delta(t-u) = (T_u\delta)(t).$$

Thus

$$x(t) = \int_{-\infty}^{\infty} x(u)(T_u\delta)(t)\,du$$

or

$$x = \int_{-\infty}^{\infty} x(u)(T_u\delta)\,du. \tag{8}$$

For any fixed u, the function

$$x_u = x(u)\,T_u\delta$$

is an impulse of magnitude $x(u)$ at time u. Thus *we have represented an arbitrary input, x, as a superposition of impulses.*

The rest is all downhill. Applying G to both sides of (8) and using the extended superposition property, (L1″), we obtain

$$Gx = \int_{-\infty}^{\infty} G[x(u)\, T_u \delta]\, du$$

$$= \int_{-\infty}^{\infty} x(u)\, G(T_u \delta)\, du,$$

by (L2), since, for fixed u, $x(u)$ is just a constant;

$$= \int_{-\infty}^{\infty} x(u)\, T_u(G\delta)\, du,$$

by invariance (property (I));

$$= \int_{-\infty}^{\infty} x(u)\, T_u\, g\, du,$$

by the definition of g. But

$$Gx = \int_{-\infty}^{\infty} x(u)\, T_u\, g\, du$$

means that

$$Gx(t) = \int_{-\infty}^{\infty} x(u)\, T_u\, g(t)\, du$$

for all t, or

$$Gx(t) = \int_{-\infty}^{\infty} g(t-u)\, x(u)\, du, \tag{9}$$

by the definition of the shift transformation.

Finally, since G is nonanticipating, and since $\delta(t) = 0$ for $t < 0$, it follows that $g(t) = G\delta(t) = 0$ for $t < 0$ (see property (N′) in the last lecture). Hence

$$Gx(t) = \int_{-\infty}^{t} g(t-u)\, x(u)\, du, \tag{1}$$

as was to be shown.

Spatial Patterns

Suppose that we are interested in the appearance of a temporally constant pattern of vertical stripes. If $x(t)$ is the luminance at horizontal position t, and $y(t) = Gx(t)$ is the apparent brightness at t, we might be tempted to consider whether G is linear, invariant, and/or nonanticipating. Linearity and invariance are interesting properties in this context. (Note that "invariance" is spatial invariance, as opposed to the temporal invariance considered above.) However, nonanticipation represents a gross left–right asymmetry that has no relation to visual perception. Hence it is of interest to know how our representation, (1), of the most general linear invariant nonanticipating system is affected if the nonanticipation condition is dropped. Our arguments show that the resulting more general representation is (9).

THEOREM 3. *Equation* (9) *represents the most general invariant linear system. The system is nonanticipating if and only if* $g(t) = G\delta(t) = 0$ *for* $t < 0$, *in which case* (9) *reduces to* (1).

For full discussions of some linear systems approaches to the perception of spatial patterns, see Harris (1980) and Chap. XII of Cornsweet (1970).

g May Be a Generalized Function

The impulse response of the *RC* circuit is the bona fide function

$$\begin{aligned} g(t) &= b^{-1}e^{-t/b} && \text{for} \quad t \geqslant 0, \\ &= 0 && \text{for} \quad t < 0, \end{aligned}$$

illustrated in Fig. 7. However, there are simple systems for which g is a close relative of δ. For example, if $Gx = 5x$, then $g = G\delta = 5\delta$. If $G = T_\tau$, then $g = T_\tau\delta$, a unit impulse at time τ. And if $Gx = dx/dt$, then $g = d\delta/dt$, whatever *that* means!

Exercise

If x has finite onset time, let $y = Gx$ satisfy the differential equation $y' = x$ and the initial condition $t_y \geqslant t_x$. Find the impulse response of G. (Hint: Integrate both sides of $y' = x$ from $-\infty$ to t, then compare the resulting expression with (1).) G is called an *integrator.*

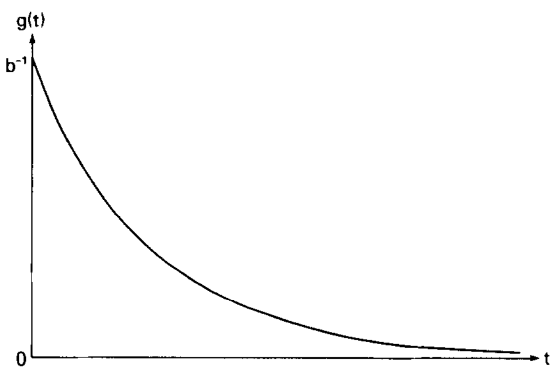

FIG. 7. Impulse response of an *RC* circuit.

Lecture 4. *Complex Numbers*

Our next main objective is to analyze an electrical circuit containing an inductance, as well as a resistance and a capacitance. Such a circuit can exhibit sinusoidal oscillation, and the study of such oscillation is enormously simplified by the introduction of complex numbers. Hence this is a good time to spend a lecture reviewing some of the elements of the theory of complex numbers. Chapter 22 of Feynman, Leighton, and Sands (1963) is required reading, and Chapter II of Knopp (1952) is also recommended.

A complex number is an expression of the form $z = x + iy$, where x and y are real numbers (not input and output functions), and $i = \sqrt{-1}$. Alternatively, we may think of z as a pair of real numbers, $z = (x, y)$. This suggests the graphical representation in the Cartesian plane that is shown in Fig. 8. This representation is sometimes called an *Argand diagram* or a *complex plane.* x and y are termed the *real* and *imaginary parts* of z, and denoted

$$x = \operatorname{re} z, \qquad y = \operatorname{im} z.$$

z is real if and only if $y = 0$.

The algebra of complex numbers involves two operations:

(1) *Addition.*

$$\begin{aligned} z + z' &= (x + iy) + (x' + iy') \\ &= (x + x') + i(y + y'). \end{aligned}$$

This means that $\operatorname{re}(z + z') = \operatorname{re} z + \operatorname{re} z'$ and $\operatorname{im}(z + z') = \operatorname{im} z + \operatorname{im} z'$.

(2) *Multiplication.*

$$\begin{aligned} zz' &= xx' + i^2yy' + ixy' + iyx' \\ &= (xx' - yy') + i(xy' + yx'). \end{aligned}$$

If $z = x$ is real, this reduces to

$$xz' = xx' + ixy'.$$

Geometrically, the sum, $z + z'$, is obtained by "completing the parallelogram," as

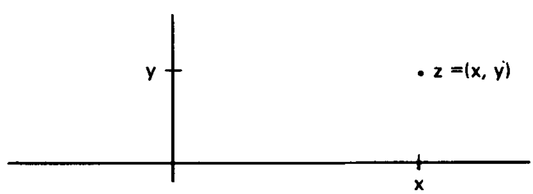

FIG. 8. The complex plane.

shown in Fig. 9. The geometric interpretation of complex multiplication will be given later.

The basic algebraic properties of complex numbers are identical to the analogous properties of real numbers. Thus $zz' = z'z$, $z(w + w') = zw + zw'$, etc.

If $z = x + iy$, its *conjugate* is

$$\bar{z} = x - iy.$$

Conjugates are illustrated in Fig. 10. Note that

$$z\bar{z} = x^2 + y^2$$

is real. The square root of this quantity is the distance, r, of z (or $\bar{z}$) from 0 in the complex plane. It is denoted $|z|$ and called the *modulus* or *absolute value* of z,

$$r = |z| = \sqrt{z\bar{z}} = \sqrt{x^2 + y^2}.$$

We can specify z completely by giving its modulus and also the angle, θ, between z and the positive real axis. See Fig. 11. The angle θ is called the *argument* of z and denoted

$$\theta = \arg z.$$

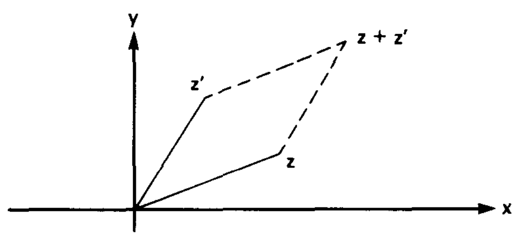

FIG. 9. Addition of complex numbers.

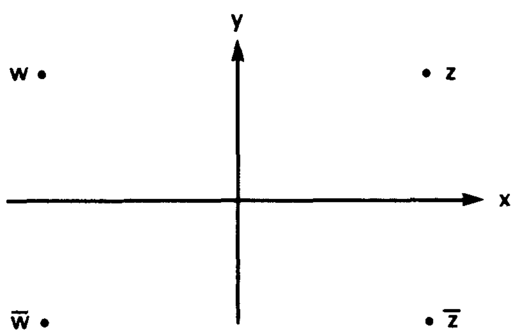

FIG. 10. Complex conjugates.

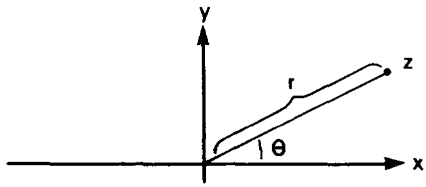

FIG. 11. Modulus, r, and argument, θ.

r and θ are the *polar coordinates* of z. In theoretical work, angles are usually measured in *radians*, rather than degrees. 2π radians $= 360$ degrees.

It is easy to express the "rectangular coordinates," x and y, in terms of the polar coordinates. By definition

$$\cos\theta = x/r, \qquad \sin\theta = y/r.$$

Hence

$$x = r\cos\theta, \qquad y = r\sin\theta. \tag{1}$$

The Exponential Function

You may have seen the representation of the real exponential function as an infinite (Taylor) series

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!},$$

where $0! = 1$ and

$$k! = k \times k-1 \times \cdots \times 2 \times 1$$

for $k \geqslant 1$. Since the only operations involved on the right are addition and multiplication, which make sense for complex numbers, it is feasible to use the series to *define* e^z for z complex,

$$e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}. \tag{2}$$

Using this definition, it is not hard to show that

$$e^{z+w} = e^z e^w, \tag{3}$$

just as for the real exponential function.

Now we come to an amazing fact, discovered by the great analyst L. Euler (1707–1783). *If θ is real, $e^{i\theta}$ has modulus* 1 *and argument θ,*

$$|e^{i\theta}| = 1, \qquad \arg e^{i\theta} = \theta.$$

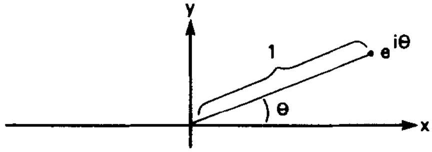

FIG. 12. The complex number $e^{i\theta}$.

This is illustrated in Fig. 12. By (1), the rectangular coordinates of $e^{i\theta}$ are

$$\operatorname{re} e^{i\theta} = \cos\theta, \qquad \operatorname{im} e^{i\theta} = \sin\theta.$$

Thus

$$e^{i\theta} = \cos\theta + i\sin\theta \tag{4}$$

("Euler's formula"). It is well to ponder this state of affairs. As θ increases, $e^{i\theta}$ executes uniform counterclockwise motion around the unit circle (circle of radius 1) in the complex plane. This motion is conceptually simple. But the motion of the projections, $\cos\theta$ and $\sin\theta$, on the real and imaginary axes is much more complicated. Thus we shall often use $e^{i\theta}$ instead of $\sin\theta$ and $\cos\theta$ to represent simple harmonic motion.

We can now rewrite (1) in an interesting way. First

$$\begin{aligned} z &= x + iy \\ &= r\cos\theta + ir\sin\theta \\ &= r(\cos\theta + i\sin\theta). \end{aligned}$$

Hence, by Euler's formula,

$$z = re^{i\theta} \qquad (\text{"polar form of } z\text{"}), \tag{5}$$

where

$$r = |z| \qquad \text{and} \qquad \theta = \arg z.$$

Polar representation is illustrated in Fig. 13.

Now $e^{i\theta}$, like $\sin\theta$ and $\cos\theta$, repeats itself when θ increases by 2π. Hence we also have

$$z = re^{i(\theta + 2\pi n)}, \qquad n = 0, \pm 1, \pm 2, \ldots.$$

In other words, the argument or angle of z is really determined only up to an additive integer multiple of 2π. The statements $\arg i = \pi/2$ and $\arg i = \pi/2 + 2\pi = 5\pi/2$ are equivalent.

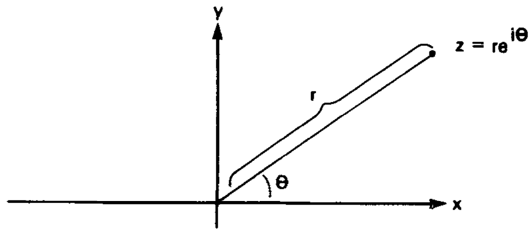

FIG. 13. The complex number $re^{i\theta}$.

We can use (5) to give a geometric interpretation of complex multiplication. Suppose that

$$z = \mathrm{r}e^{i\theta} \qquad \text{and} \qquad z' = r'e^{i\theta'}.$$

Then

$$\begin{aligned} zz' &= rr'e^{i\theta}e^{i\theta'} \\ &= rr'e^{i(\theta+\theta')}. \end{aligned}$$

Thus

$$|zz'| = rr' = |z|\,|z'|$$

and

$$\arg(zz') = \theta + \theta' = \arg z + \arg z'.$$

Hence the modulus of a product is the product of the moduli and the angle of a product is the sum of the angles. For example, $|i| = 1$ and $\arg i = \pi/2$, so

$$|iz| = |z|,\ \arg(iz) = \pi/2 + \arg z.$$

In other words, multiplication by i causes counterclockwise rotation by a right angle. The formula for $\arg(zz')$ also implies that

$$\arg z^n = n \arg z$$

for any positive integer n.

In subsequent lectures, we are going to be much concerned with the *complex sinusoid*

$$z(t) = e^{st},$$

with *complex frequency* $s = \sigma + i\omega$. This is a complex-valued function of the real variable t, which, in these lectures, ususally represents time. (In the brightness perception problem mentioned near the end of Lecture 3, t represents horizontal position.) Clearly

$$z(t) = e^{(\sigma + i\omega)t}$$
$$= e^{\sigma t + i\omega t}$$

or

$$z(t) = e^{\sigma t} e^{i\omega t}. \tag{6}$$

This is the polar from of $z(t)$. The modulus and argument of $z(t)$ are thus

$$r(t) = |z(t)| = e^{\sigma t}$$

and

$$\theta(t) = \arg z(t) = \omega t.$$

As t increases, $z(t)$ rotates counterclockwise at a constant angular velocity, ω,

$$d\theta/dt = \omega = \text{angular velocity},$$

while the modulus of $z(t)$ increases or decreases exponentially, depending on whether $\sigma > 0$ or $\sigma < 0$. Hence $z(t)$ spirals outward if $\sigma > 0$ and $z(t)$ spirals inward if $\sigma < 0$. If $\sigma = 0$ then $z(t) = e^{i\omega t}$, which runs around the unit circle at ω radians per unit time.

The rectangular coordinates of $z(t)$ are also interesting. Equation (6) implies that

$$z(t) = e^{\sigma t} \cos \omega t + i e^{\sigma t} \sin \omega t,$$

so

$$\text{re } e^{st} = e^{\sigma t} \cos \omega t = x(t)$$

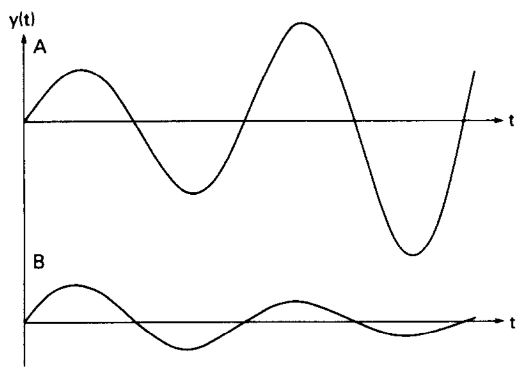

FIG. 14. Exponentially modulated sinusoids, $y(t) = e^{\sigma t} \sin \omega t$. In Panel A, $\sigma > 0$. In Panel B, $\sigma < 0$.

and

$$\text{im } e^{st} = e^{\sigma t} \sin \omega t = y(t).$$

These are amplitude modulated sinusoids. Figure 14 illustrates $y(t)$ for positive and negative σ.

Differentiation

The standard methods of calculus extend easily to complex-valued functions of real variables. Let $z(t) = x(t) + iy(t)$ be *any* such function. We define the derivative of z just as we would if z were real valued,

$$z'(t) = \frac{dz}{dt} = \lim_{\delta \to 0} \delta^{-1}[z(t+\delta) - z(t)], \tag{7}$$

and we note that it follows that

$$\begin{aligned} z'(t) &= \lim_{\delta \to 0} \delta^{-1}[x(t+\delta) + iy(t+\delta) - x(t) - iy(t)] \\ &= \lim_{\delta \to 0} \delta^{-1}[x(t+\delta) - x(t) + i(y(t+\delta) - y(t))] \\ &= \lim_{\delta \to 0} [\delta^{-1}(x(t+\delta) - x(t)) + i\delta^{-1}(y(t+\delta) - y(t))]. \end{aligned}$$

Thus

$$z'(t) = x'(t) + iy'(t). \tag{8}$$

Therefore differentiation of a complex-valued function of a real variable is tantamount to differentiating its real and imaginary parts. (Note that z' in this section is the derivative of the complex function z, whereas, in earlier sections, z' was just a complex number that I wished to distinguish from another complex number, z.)

For real s, we know that

$$\frac{d}{dt} e^{st} = se^{st}. \tag{9}$$

We shall now show that this important formula is also valid for complex s. In view of (7),

$$\begin{aligned} \frac{d}{dt} e^{st} &= \lim_{\delta \to 0} \delta^{-1}[e^{s(t+\delta)} - e^{st}] \\ &= \lim_{\delta \to 0} \delta^{-1}[e^{s\delta} - 1]\, e^{st}, \end{aligned}$$

by (3);

$$= \lim_{\delta \to 0} \delta^{-1} \left[s\delta + \frac{s^2\delta^2}{2!} + \frac{s^3\delta^3}{3!} + \cdots \right] e^{st},$$

by (2);

$$= \lim_{\delta \to 0} \left[s + \frac{s^2\delta}{2!} + \frac{s^3\delta^2}{3!} + \cdots \right] e^{st}$$

$$= se^{st},$$

as was to be shown.

In the special case $s = i$, (9) reduces to

$$\frac{d}{dt} e^{it} = ie^{it}.$$

Using (8) on the left, this becomes

$$\frac{d}{dt} \cos t + i \frac{d}{dt} \sin t = i(\cos t + i \sin t) = -\sin t + i \cos t.$$

But equality of two complex numbers is equivalent to equality of their real and imaginary parts, so we have obtained the famous formulas

$$\frac{d}{dt} \sin t = \cos t$$

and

$$\frac{d}{dt} \cos t = -\sin t,$$

which have mystified students in elementary calculus courses for decades.

Exercises

Study the lecture before beginning the exercises. Solutions to the exercises should be based on material presented in the lecture.

1. Show that $e^0 = 1$. (Hint: Consider the definition of e^z.)
2. Prove that $e^{-w} = (e^w)^{-1}$. (Hint: Consider $e^{-w}e^w$.)
3. If $z = re^{i\theta}$, what are the polar coordinates of z^{-1}? (Hint: $(vw)^{-1} = v^{-1}w^{-1}$.)
4. If $z = x + iy$, what are the rectangular coordinates of $1/z$? (Hint: Multiply numerator and denominator by $\bar{z}$.)

5. If $z = re^{i\theta}$, show graphically that $\bar{z} = re^{-i\theta}$; in particular,

$$\overline{e^{i\theta}} = e^{-i\theta}. \tag{10}$$

6. Prove that $\cos(-\theta) = \cos\theta$ and $\sin(-\theta) = -\sin\theta$. (Hint: Express both sides of (10) in rectangular coordinates.)

7. Prove that $\overline{zw} = \bar{z}\bar{w}$. (Hint: Use polar representations of z and w.)

8. Prove that $\operatorname{re} z = (z + \bar{z})/2$ and $\operatorname{im} z = (z - \bar{z})/(2i)$. Use the latter result to prove the obvious fact that z is real if and only if $z = \bar{z}$.

9. Prove that $\cos\theta = (e^{i\theta} + e^{-i\theta})/2$ and $\sin\theta = (e^{i\theta} - e^{-i\theta})/(2i)$. (Hint: Remember (10).)

10. Show graphically that $e^{i\pi/2} = i$ and $e^{i\pi} = -1$. Use the latter result and Exercise 9 to prove that $\cos(\theta + \pi) = -\cos\theta$ and $\sin(\theta + \pi) = -\sin\theta$.

11. Prove that

$$\cos(\theta_1 + \theta_2) = \cos\theta_1 \cos\theta_2 - \sin\theta_1 \sin\theta_2,$$
$$\sin(\theta_1 + \theta_2) = \sin\theta_1 \cos\theta_2 + \cos\theta_1 \sin\theta_2.$$

(Hint: Express both sides of $e^{i(\theta_1 + \theta_2)} = e^{i\theta_1}e^{i\theta_2}$ in rectangular coordinates.)

12. Prove that $|z + w| \leqslant |z| + |w|$. (A graphical argument will suffice.)

13. "Verify" $e^{i\pi} = -1$ by calculating $\sum_{k=0}^{10} (i\pi)^k/k!$.

Lecture 5. RCL Circuits

In Lecture 1 we analyzed an electrical circuit with a resistor and a capacitor in series. This led to a differential equation relating $y(t)$, the voltage across the capacitor, and $y'(t)$ to $x(t)$, the impressed voltage. Since the equation involves only the first derivative, it is called a *first order equation.* If we introduce an inductance into the circuit, the resulting differential equation involves y'', and is hence a *second order equation.* Figure 15 illustrates a circuit containing a resistor, a capacitor, and an inductance in series.

Arguing just as we did in the first lecture, we obtain

$$v_{14} = v_{12} + v_{23} + v_{34}, \tag{1}$$

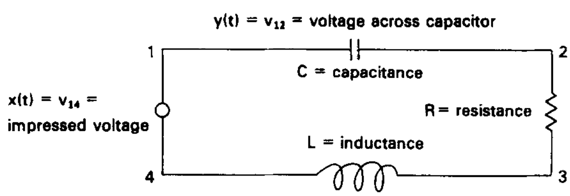

FIG. 15. An *RCL* circuit.

where

$$v_{14} = x(t),$$
$$v_{12} = y(t),$$
$$v_{23} = i(t)R = y'(t)\,RC,$$
$$v_{34} = i'(t)L = y''(t)\,LC.$$

Only the equation for the voltage, v_{34}, across the inductance is new. It expresses the fact that an inductance opposes a change in current, just as a mass opposes a change in its velocity. If current increases, a voltage is created that opposes the flow of electrons. Substituting into (1), we find that

$$ay'' + by' + y = x, \tag{2}$$

where

$$a = LC \qquad \text{and} \qquad b = RC.$$

Since $y = C^{-1}q$, where q is the charge on the capacitor, (2) can be expressed in terms of q as follows

$$Lq'' + Rq' + C^{-1}q = x. \tag{3}$$

An analogous mechanical system is a mass suspended by a spring in a container of viscous fluid, as shown in Fig. 16. If y is the displacement of the mass from its resting position, and if x is an external force on the mass (over and above the gravitational force) then

$$my'' + cy' + ky = x, \tag{4}$$

by elementary mechanics. The term cy' represents the force on the mass due to its motion through the fluid. This force is proportional to its velocity, at least if velocity

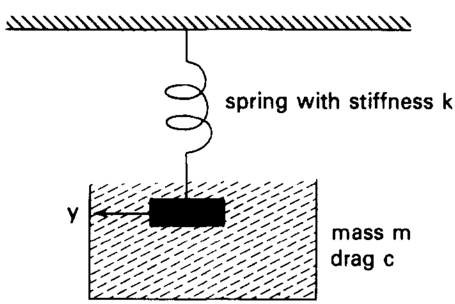

FIG. 16. A simple mechanical system.

is not too large. The drag coefficient, c, depends on the viscosity of the fluid, among other things. Comparing (3) and (4), we obtain the dictionary

electrical circuit	charge, q	inductance, L	resistance, R	C^{-1}
mechanical rig	displacement, y	mass, m	drag, c	stiffness, k

Thus, in the process of studying the electrical circuit, we are also obtaining the behavior of the mechanical rig, and of a variety of other important systems that arise in electricity, mechanics, and other domains.

We now return to the differential equation, (2), for the voltage, y, across the capacitor of the RCL circuit. For the time being, we shall restrict our attention to inputs, x, whose onset times satisfy $t_x > -\infty$. Thus t_x is finite, or $x(t) = 0$ for all t, in which case $t_x = +\infty$. It can be shown that, *if* $t_x > -\infty$, *then* (2) *has one and only one solution*, y, *for which* $t_y > -\infty$. *In fact, as you would expect*, $t_y \geqslant t_x$. *We define a system*, G, *by letting* Gx *be this solution*; *thus* $Gx = y$.

THEOREM 1. *G is linear, invariant, and nonanticipating.*

Proof. The condition $t_y \geqslant t_x$ is called (N′) in Lecture 2, where it is noted that this condition is equivalent to nonanticipation for linear systems. Thus we need only prove linearity and invariance.

Let x_1 and x_2 be inputs with $t_{x_1} > -\infty$ and $t_{x_2} > -\infty$, and let $y_1 = Gx_1$ and $y_2 = Gx_2$ be the corresponding outputs. Then

$$ay_1'' + by_1' + y_1 = x_1$$

and

$$ay_2'' + by_2' + y_2 = x_2.$$

Adding these equations, we get

$$a(y_1 + y_2)'' + b(y_1 + y_2)' + (y_1 + y_2) = x_1 + x_2, \tag{5}$$

since $y_1' + y_2' = (y_1 + y_2)'$ and $y_1'' + y_2'' = (y_1 + y_2)''$ (differentiation is linear!). According to (5),

$$ay'' + by' + y = x,$$

where

$$y = y_1 + y_2 \qquad \text{and} \qquad x = x_1 + x_2.$$

Moreover, $t_y > -\infty$, since $t_{y_1} > -\infty$ and $t_{y_2} > -\infty$. Thus

$$y = Gx$$

or

$$y_1 + y_2 = G(x_1 + x_2),$$

or

$$Gx_1 + Gx_2 = G(x_1 + x_2).$$

Hence G has the superposition property, (L1). A similar argument verifies (L2).

To establish invariance, apply the shift operator, T_τ, to both sides of (2). This yields

$$a(T_\tau y)'' + b(T_\tau y)' + T_\tau y = T_\tau x,$$

since $T_\tau(y') = (T_\tau y)'$ and $T_\tau(y'') = (T_\tau y)''$ (differentiation is invariant). Thus

$$aY'' + bY' + Y = X,$$

where

$$Y = T_\tau y \qquad \text{and} \qquad X = T_\tau x.$$

Clearly $t_Y > -\infty$, so

$$Y = GX$$

or

$$T_\tau y = GT_\tau x$$

or

$$T_\tau Gx = GT_\tau x.$$

Thus G is invariant.

This completes the proof of the theorem. Similar arguments apply to any system described by a differential equation with constant coefficients. These arguments are based on the fact that the differential equation has one and only one solution satisfying the initial condition $t_y > -\infty$. This is an "existence and uniqueness theorem" for the differential equation. Such theorems are readily available in mathematics books, but they require some work to prove.

Impulse Response

Now that we know that G is an invariant linear system, we may, in good conscience, proceed to determine its impulse response, $g = G\delta$. We expect this function to be the solution of

$$ag''(t) + bg'(t) + g(t) = \delta(t) \tag{6}$$

with $t_g \geqslant t_\delta = 0$. The meaning of the differential equation is a bit murky because of the delta "function" on the right, but we shall be able to extract enough information from (6) to determine g.

First, we note that $\delta(t) = 0$ for $t > 0$. Thus, for $t > 0$, (6) reduces to

$$ag''(t) + bg'(t) + g(t) = 0. \tag{7}$$

As we shall see momentarily, this differential equation has an infinite number of solutions. To select the impulse response from among these solutions, we need additional information about the impulse response. The necessary information is contained in the two initial values

$$g(0+) = \lim_{t \downarrow 0} g(t)$$

and

$$g'(0+) = \lim_{t \downarrow 0} g'(t),$$

which pertain to the behavior of the system immediately after an impulse input.

How about $g(0+)$? If we "sock" the RCL circuit with an impulse at time 0, does voltage instantaneously appear across the capacitor? It does in the RC circuit ($g(t) = b^{-1}e^{-t/b}$ for $t > 0$, so $g(0+) = b^{-1}$). However, it doesn't here, since the inductance "fights" the surge of current and thus retards the charging of the capacitor. The situation is even clearer with the mechanical rig. If you kick the mass, it doesn't "jump." Instead, it "takes off." Thus $g(0+) = 0$ but $g'(0+) > 0$.

What, precisely, is the initial rate of change, $g'(0+)$, of the output voltage, due to an impulsive input at time 0? To determine this, we return to (6) and integrate from $t = -\varepsilon$ to $t = \varepsilon$. This yields

$$a(g'(\varepsilon) - g'(-\varepsilon)) + b(g(\varepsilon) - g(-\varepsilon)) + \int_{-\epsilon}^{\epsilon} g(t)\, dt$$

$$= \int_{-\epsilon}^{\epsilon} \delta(t)\, dt = 1.$$

Moreover, $g(t) = 0$ for all $t < 0$, so $g(-\varepsilon) = 0$, $g'(-\varepsilon) = 0$, and

$$\int_{-\epsilon}^{\epsilon} g(t)\, dt = \int_{0}^{\epsilon} g(t)\, dt.$$

Thus

$$ag'(\varepsilon) + bg(\varepsilon) + \int_{0}^{\epsilon} g(t)\, dt = 1. \tag{8}$$

Now let $\varepsilon \downarrow 0$. We know that $g(\varepsilon) \to g(0+) = 0$, and, clearly,

$$\int_0^\varepsilon g(t)\, dt \to \int_0^0 g(t)\, dt = 0.$$

Hence (8) yields, in the limit, $g'(0+) = a^{-1}$.

In summary, we have found that $g = G\delta$ satisfies

$$ag''(t) + bg'(t) + g(t) = 0 \tag{7}$$

for $t > 0$, together with the initial conditions

$$g(0+) = 0 \tag{9}$$

and

$$g'(0+) = a^{-1}. \tag{10}$$

It remains only to show that this information suffices to determine g.

Putting aside initial conditions for the moment, let us determine solutions of the differential equation, (7). Our complex sinusoid e^{st} comes in handy here. We know that

$$de^{st}/dt = se^{st},$$

hence

$$d^2e^{st}/dt^2 = s^2e^{st}.$$

Thus e^{st} is a solution of (7) if and only if

$$as^2e^{st} + bse^{st} + e^{st} = 0$$

or

$$(as^2 + bs + 1)\, e^{st} = 0. \tag{11}$$

Now $e^{st} \neq 0$ (since $e^{st}e^{-st} = e^0 = 1$), so the product in (11) is zero if and only if the first factor is zero. Thus e^{st} is a solution of (7) if and only if s satisfies the *characteristic equation*

$$as^2 + bs + 1 = 0.$$

This equation has precisely two solutions,

$$s_+ = \frac{-b + \sqrt{d}}{2a}$$

and

$$s_- = \frac{-b - \sqrt{d}}{2a},$$

where

$$d = b^2 - 4a.$$

These expressions represent the famous "quadratic formula," and d is called the *discriminant*. We conclude that $e^{s_+ t}$ and $e^{s_- t}$ are the only complex sinusoids that solve (7).

These solutions can be combined to obtain additional solutions. Arguments presented earlier in this lecture show that, for any constants c_+ and c_-, $c_+ e^{s_+ t}$ and $c_- e^{s_- t}$ are solutions of (7), as is

$$c_+ e^{s_+ t} + c_- e^{s_- t}.$$

Moreover, it can be shown that this is the most general solution of (7), provided that $s_+ \neq s_-$ (or, equivalently, $d \neq 0$). In other words, if $s_+ \neq s_-$ and if g is any solution of (7), there are constants, c_+ and c_-, such that

$$g(t) = c_+ e^{s_+ t} + c_- e^{s_- t}. \tag{12}$$

In particular, the impulse response $g = G\delta$ has this form for $t > 0$.

The constants c_+ and c_- are determined by the initial conditions (9) and (10). Letting $t \downarrow 0$ in (12), we obtain

$$g(0+) = c_+ + c_-.$$

Thus (9) is equivalent to

$$c_+ + c_- = 0. \tag{9$'$}$$

Differentiating (12) and then letting $t \downarrow 0$, we get

$$g'(0+) = c_+ s_+ + c_- s_-,$$

so (10) is equivalent to

$$c_+ s_+ + c_- s_- = a^{-1}. \tag{10$'$}$$

It is easy to solve the simultaneous linear equations (9′) and (10′) for c_+ and c_-. The first equation is equivalent to $c_- = -c_+$. Replacing c_- in (10′) by $-c_+$, we obtain

$$c_+(s_+ - s_-) = a^{-1}$$

or

$$c_+ = \frac{1}{a(s_+ - s_-)} = \frac{1}{\sqrt{d}}.$$

Thus

$$c_- = -c_+ = -\frac{1}{\sqrt{d}}.$$

Combining these expressions with (12), we find that $g = G\delta$ is given by

$$g(t) = \frac{1}{\sqrt{d}}(e^{s_+t} - e^{s_-t}) \tag{13}$$

for $t > 0$, provided that $d = b^2 - 4a \neq 0$. There is no harm in also using (13) for $t = 0$. This yields $g(0) = 0 = g(0+)$. Since the RCL circuit is nonanticipating, $g(t) = 0$ for $t < 0$. The following theorem summarizes these results.

THEOREM 2. *Let $a = LC$, $b = RC$, $d = b^2 - 4a$, and*

$$s_\pm = \frac{-b \pm \sqrt{d}}{2a}.$$

If $d \neq 0$, the impulse response of the RCL circuit is $g(t) = 0$ for $t \leqslant 0$ and

$$g(t) = \frac{1}{\sqrt{d}}(e^{s_+t} - e^{s_-t}) \tag{13}$$

for $t \geqslant 0$.

The comparable result for $d = 0$ is as follows.

THEOREM 3. *If $d = 0$, then*

$$g(t) = a^{-1}te^{-\sigma t} \tag{14}$$

for $t \geqslant 0$, where $\sigma = b/(2a)$.

A derivation of (14) can be given along the lines of our derivation of (13). We leave this as an exercise.

Exercises

Note that, if $d = 0$, then $s_+ = s_- = -\sigma$, so $e^{-\sigma t}$ is the only complex exponential solution of (7). However, $te^{-\sigma t}$ is also a solution in this case.

1. Verify that $te^{-\sigma t}$ satisfies (7) if $d = 0$.

It can be shown that the most general solution of (7) is

$$g(t) = k_1 e^{-\sigma t} + k_2 t e^{-\sigma t}.$$

2. Show that this function satisfies the initial conditions (9) and (10) if and only if $k_1 = 0$ and $k_2 = a^{-1}$.

Lecture 6. RCL Circuits, Convolution, Simple Composite Systems, Odds and Ends

Our next project is to take a closer look at the impulse response of the *RCL* circuit. Recall that, if $d = 0$, then

$$g(t) = a^{-1} t e^{-\sigma t} \tag{1}$$

for $t \geqslant 0$, whereas if $d \neq 0$, then

$$g(t) = \frac{1}{\sqrt{d}} (e^{s_+ t} - e^{s_- t}) \tag{2}$$

for $t \geqslant 0$, where

$$\begin{aligned} \sigma &= b/(2a), \\ s_{\pm} &= \frac{-b \pm \sqrt{d}}{2a}, \end{aligned} \tag{3}$$

and

$$d = b^2 - 4a.$$

If $d \neq 0$, it may be positive or negative. Let us consider each of these possibilities.

Suppose first that $d > 0$. Then s_+ and s_- are both real, and $s_+ > s_-$. Moreover $a = LC > 0$, so

$$d = b^2 - 4a < b^2.$$

Hence

$$\sqrt{d} < b,$$

so that

$$\begin{aligned} s_+ &= \frac{-b + \sqrt{d}}{2a} \\ &< \frac{-b + b}{2a} \\ &= 0. \end{aligned}$$

Therefore, if $d > 0$,

$$s_- < s_+ < 0.$$

It follows that $s_- t < s_+ t$ and $e^{s_- t} < e^{s_+ t}$ if $t > 0$. Therefore $g(t) > 0$ for all $t > 0$. Also $s_\pm t \to -\infty$ as $t \to \infty$, so $e^{s_\pm t} \to 0$ as $t \to \infty$. Consequently $g(t) \to 0$ as $t \to \infty$.

The simplest picture consistent with these observations and with $g(0) = 0$ is shown in Fig. 17. The distinguishing feature of this graph is the single peak. It is not difficult to show that the graph of g does indeed have this shape, and that the peak occurs at

$$t = (s_+ - s_-)^{-1} \ln(s_-/s_+),$$

where ln is the natural logarithm. (This expression for t was obtained by differentiating (2), setting the derivative equal to zero, and solving the resulting equation.)

The shape of the impulse response for $d = 0$ is similar to that shown in Fig. 17. However the graph is rather different when $d < 0$, as we shall now show.

If $d < 0$, then $d = -|d|$, so

$$\begin{aligned} \sqrt{d} &= \sqrt{-|d|} \\ &= \sqrt{-1}\,\sqrt{|d|} \\ &= i\sqrt{|d|}. \end{aligned}$$

Thus s_+ and s_- are complex conjugates,

$$s_\pm = \frac{-b \pm i\sqrt{|d|}}{2a}$$

or

$$s_\pm = -\sigma \pm i\omega,$$

where

$$\sigma = b/2a$$

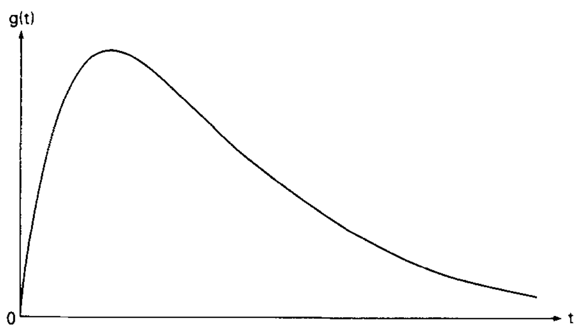

FIG. 17. Impulse response of an RCL circuit with $d > 0$.

and

$$\omega = \sqrt{|d|}/2a.$$

Substituting $-\sigma \pm i\omega$ for $s_\pm$ in (2), factoring out $e^{-\sigma t}$, and recalling that

$$\sqrt{d} = i\sqrt{|d|} = 2ia\omega,$$

we obtain

$$g(t) = \frac{1}{a\omega} e^{-\sigma t} \frac{1}{2i} (e^{i\omega t} - e^{-i\omega t}).$$

But

$$\frac{1}{2i} (e^{i\omega t} - e^{-i\omega t}) = \sin \omega t$$

(see Exercise 9 of Lecture 4). Hence, if $d < 0$,

$$g(t) = (a\omega)^{-1} e^{-\sigma t} \sin \omega t \tag{4}$$

for $t \geqslant 0$, where $\sigma = b/(2a)$ and $\omega = \sqrt{|d|}/(2a)$. This function is illustrated in Panel B of Fig. 14.

Unlike the impulse response function for $d \geqslant 0$ (or for the RC circuit), this one oscillates, crossing the t axis at points where $\omega t = n\pi$, $n = 0, 1, 2,\ldots$; that is, $t = n\pi/\omega$. The parameter ω is the frequency of free oscillation of the system. It is called the *natural frequency*.

Convolution

We showed in Lecture 3 that every invariant linear system is related to its impulse response via the equation

$$Gx(t) = \int_{-\infty}^{\infty} g(t-u)\, x(u)\, du. \tag{5}$$

If G is nonanticipating, then $g(t-u) = 0$ for $u > t$, so (5) reduces to

$$Gx(t) = \int_{-\infty}^{t} g(t-u)\, x(u)\, du.$$

However, the more general expression, (5), is more convenient for our present purposes.

Introducing the change of variables $\tau = t - u$ into (5), and noting that $u = t - \tau$, we obtain

$$Gx(t) = \int_{-\infty}^{\infty} g(\tau)\, x(t-\tau)\, d\tau.$$

Reversing the order of $g(\tau)$ and $x(t-\tau)$ and changing the name of the variable of integration from τ to u, we find that

$$Gx(t)=\int_{-\infty}^{\infty} x(t-u)\, g(u)\, du. \tag{6}$$

Notice that the roles of g and x are precisely reversed in the integrals (5) and (6).

We shall now introduce a useful notation. The integral on the right in (5) is denoted $g * x(t)$,

$$g * x(t)=\int_{-\infty}^{\infty} g(t-u)\, x(u)\, du. \tag{7}$$

For any two functions g and x, $g * x$ is a function, called the *convolution* of g and x. The value of this function at time t, $g * x(t)$, is given by (7).

In this notation, the integral on the right in (6) is $x * g(t)$, and the fact that the integrals in (5) and (6) are equal means that $g * x(t)=x * g(t)$ for all t, or

$$g * x = x * g. \tag{8}$$

Convolution of functions is analogous to multiplication of numbers, and (8) says that convolution, like multiplication, is commutative.

Using this new notation we can write (5) and (6) in the abbreviated forms

$$Gx = g * x \tag{5}$$

and

$$Gx = x * g. \tag{6}$$

As a simple application of (6), let us consider the case where x is the *unit step function*,

$$\begin{aligned} \mu(t) &= 1 \qquad \text{if} \quad t \geqslant 0, \\ &= 0 \qquad \text{if} \quad t < 0. \end{aligned}$$

Then $Gx = G\mu$ is called the *step response*. According to (6),

$$G\mu(t)=\int_{-\infty}^{\infty} \mu(t-u)\, g(u)\, du.$$

But $\mu(t-u)=0$ if $u>t$ and $\mu(t-u)=1$ if $u \leqslant t$, so

$$G\mu(t)=\int_{-\infty}^{t} g(u)\, du.$$

(If G is nonanticipating, this reduces to

$$G\mu(t) = \int_0^t g(u)\,du.) \tag{9}$$

Thus *the step response is the integral of the impulse response.* This implies that

$$g(t) = \frac{dG\mu(t)}{dt}. \tag{10}$$

The impulse response is the derivative of the step response.

Simple Composite Systems

A popular type of model in psychology and many other sciences is a network of interconnected simple components (delays, integrators, differentiators, RC circuits, etc.). The analysis of networks is greatly facilitated by frequency domain techniques treated in subsequent lectures. All I will do today is consider two simple composite systems that can be obtained by interconnecting invariant linear components, G_1 and G_2.

The composite systems I have in mind are denoted $G_1 + G_2$ and $G_1 G_2$, and defined as follows:

$$(G_1 + G_2)x = G_1 x + G_2 x \tag{11}$$

and

$$(G_1 G_2)x = G_1(G_2 x). \tag{12}$$

In the first case, we add the outputs of G_1 and G_2 for the same input, x. In the second case, the output of G_2 is taken as the input of G_1. In $G_1 + G_2$ the components operate *in parallel*; in $G_1 G_2$ they are *in series* or *cascade.* A trivial special case of (12) is aG_2, where a is a constant amplification factor,

$$(aG_2)x = a(G_2 x). \tag{13}$$

The *block diagrams* for (11), (12), and (13) are shown in Fig. 18.

It is easy to express the impulse responses of these systems in terms of the impulse responses, g_1 and g_2, of the components. Clearly

$$\begin{aligned}(G_1 + G_2)\delta &= G_1\delta + G_2\delta \qquad (14)\\ &= g_1 + g_2,\end{aligned}$$

$$\begin{aligned}(aG_2)\delta &= a(G_2\delta) \qquad (15)\\ &= ag_2,\end{aligned}$$

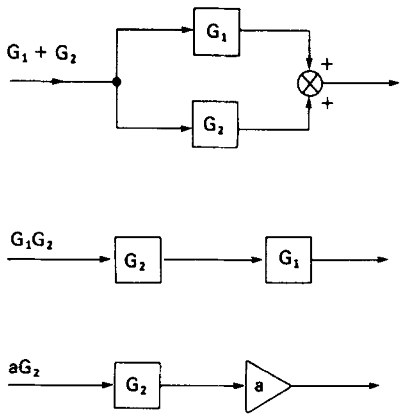

FIG. 18. Three simple composite systems.

and

$$(G_1 G_2)\delta = G_1(G_2\delta)$$
$$= G_1 g_2$$

or

$$(G_1 G_2)\delta = g_1 * g_2. \tag{16}$$

Thus the impulse response of a cascade is the corresponding convolution of impulse responses. Since $g_1 * g_2 = g_2 * g_1$ (take $g = g_1$ and $x = g_2$ in (8)), $G_1 G_2$ and $G_2 G_1$ have the same impulse response and are therefore functionally identical. *The order of components in a cascade of invariant linear systems is immaterial.* In particular, if a psychological system is a cascade of invariant linear subsystems, the order of the pieces can't be discovered by pure stimulus–response analysis. Behaviorists, take note!

Odds and Ends

Before pushing on to frequency domain analysis, let me deal with three technicalities concerning the general theory of linear systems.

1. *On linearity.* We say that G is linear if it has properties (L1), $G(x_1 + x_2) = Gx_1 + Gx_2$, and (L2), $G(ax) = aGx$. These two equations can be combined into a single equivalent equation,

$$G(a_1 x_1 + a_2 x_2) = a_1 G x_1 + a_2 G x_2. \tag{L}$$

Proof of equivalence. Suppose G satisfies (L1) and (L2). Then

$$G(a_1x_1 + a_2x_2) = G(a_1x_1) + G(a_2x_2),$$

by (L1);

$$= a_1Gx_1 + a_2Gx_2,$$

by (L2). Conversely, if G satisfies (L), then, taking $a_1 = a_2 = 1$, we see that it satisfies (L1), and, taking $a_2 = 0$, we see that it satisfies (L2).

2. *The superposition property* (L1) *almost implies homogeneity* (L2). More precisely, (L1) implies $G(ax) = aGx$ for all rational numbers a. Thus superposition is almost synonymous with linearity.

Proof. This proof is lots of fun. First, we saw in Lecture 3 that the superposition property for sums of two inputs implies an analogous property for sums of n inputs. Thus

$$\begin{aligned} G(nx) &= G(\underbrace{x + x + \cdots + x}_{n \text{ times}}) \\ &= \overbrace{Gx + Gx + \cdots + Gx} \\ &= nGx, \end{aligned}$$

for any positive integer, n.

For any positive rational number, $a = m/n$, it follows that

$$\begin{aligned} nG\left(\frac{m}{n}x\right) &= G\left(n\left(\frac{m}{n}x\right)\right) \\ &= G(mx) \\ &= mGx. \end{aligned}$$

Hence

$$G(ax) = aGx \tag{17}$$

for any positive rational number a.

Let 0 be the *null input*, $0(t) = 0$ for all t. Clearly $0 + 0 = 0$, so

$$G0 = G(0 + 0) = G0 + G0.$$

Subtracting $G0$ from both sides, we see that

$$G0 = 0. \tag{18}$$

This conclusion is of some interest in its own right. It says that any system satisfying (L1) gives no output if you provide no input. This means, in practice, that, if you wish to apply linear systems theory to a real-world system, you must let the "output," $y(t)$, be the departure from the "resting state" of the system.

Returning to our proof, (18) implies (17) for $a = 0$. It also implies

$$\begin{aligned} 0 = G0 &= G(-x + x) \\ &= G(-x) + Gx, \end{aligned}$$

hence

$$G(-x) = -Gx. \tag{19}$$

In particular, if a is a negative rational number, then $a = -|a|$, so

$$\begin{aligned} G(ax) &= G(-|a|x) \\ &= -G(|a|\,x), \end{aligned}$$

by (19);

$$= -|a|\,Gx,$$

since $|a|$ is a positive rational number;

$$= aGx.$$

Thus $G(ax) = aGx$ for all rational numbers—positive, negative, and zero—as was to be shown.

3. *If G has the superposition property, then the nonanticipation condition,*

$$x_1(t) = x_2(t) \text{ for } t < T \text{ implies } y_1(t) = y_2(t) \text{ for } t < T, \tag{N}$$

is equivalent to

$$t_y \geqslant t_x. \tag{N$'$}$$

Here $y_i = Gx_i$, $y = Gx$, and t_x is the onset time of x.

Proof. Suppose G is nonanticipating. For $t < t_x$, $x(t) = 0(t)$, where 0 is the null input. Hence (N) implies $Gx(t) = G0(t)$ for $t < t_x$. But $Gx = y$ and $G0 = 0$, so $y(t) = 0$ for $t < t_x$. Therefore $t_y \geqslant t_x$ and (N$'$) is satisfied.

Suppose, conversely, that (N$'$) holds, and that $x_1(t) = x_2(t)$ for $t < T$. Then $t_x \geqslant T$, where $x = x_1 - x_2$, so, by (N$'$), $t_y \geqslant T$, where

$$\begin{aligned} y &= Gx \\ &= Gx_1 - Gx_2 \\ &= y_1 - y_2. \end{aligned}$$

Hence $y_1(t) - y_2(t) = 0$ for $t < T$, and G is nonanticipating.

Exercises

Use (16) to find the impulse response of a cascade of two *RC* circuits, G_1 and G_2. Let b_i be the time constant of G_i. You will have to treat the cases $b_1 = b_2$ and $b_1 \neq b_2$ separately. Compare your results with Eqs. (1) and (2) and discuss the implications of this comparison for "gross behaviorism."

Part B. Analysis in the Frequency Domain

For additional information on this topic, see Bracewell (1978) or Chapters 5 and 6 of Schwarz and Friedland (1965).

Lecture 7. Transfer Functions and Laplace Transforms

Much of linear systems theory revolves around the fact that linear systems map sinusoidal inputs into sinusoidal outputs. We will prove this today. The part of the theory that deals with these matters is called "analysis in the frequency domain." Now it is possible to develop this material in terms of sines and cosines, but, as you might guess, it is much easier to use complex sinusoids, and we will follow the latter route. Thus we will consider the output, Gz, when z is the complex sinusoid $z(t) = e^{st}$. First, however, we must ask what it means to put a complex input into a real system! It means just this: A complex input,

$$z(t) = x(t) + iy(t),$$

may be thought of as a pair of real inputs. We apply G separately to each of these to obtain two outputs, Gx and Gy, which are taken to be the two components of the complex output, Gz. Definition 1 expresses this in other words.

Definition 1. If $z = x + iy$, then Gz is defined by

$$Gz(t) = Gx(t) + iGy(t)$$

or

$$Gz = Gx + iGy.$$

Think of running the "signals" x and y through separate channels. G operates independently in both channels to produce a two-channel output.

It is easy to show that, if we begin with a system that is linear, or invariant, or nonanticipating for real inputs, then the extended system defined above has the same properties for complex inputs. In the case of property (L2), if $G(ax) = aGx$ for a and x real, then the extended system satisfies $G(cz) = cGz$ for c and z complex.

We shall also need to know how to integrate complex-valued functions of a real variable.

DEFINITION 2. If $z(t) = x(t) + iy(t)$, then $\int_a^b z(t)\,dt$ is defined by

$$\int_a^b z(t)\,dt = \int_a^b x(t)\,dt + i\int_a^b y(t)\,dt.$$

Suppose now that G is the invariant linear system given by

$$Gx(t) = g * x(t) = \int_{-\infty}^{\infty} g(t-u)\,x(u)\,du, \tag{1}$$

for x real, and let $z(t) = x(t) + iy(t)$ be a complex input. Then

$$Gz(t) = Gx(t) + iGy(t),$$

by Definition 1;

$$= \int_{-\infty}^{\infty} g(t-u)\,x(u)\,du + i\int_{-\infty}^{\infty} g(t-u)\,y(u)\,du,$$

by Eq. (1);

$$= \int_{-\infty}^{\infty} [g(t-u)\,x(u) + ig(t-u)\,y(u)]\,du,$$

by Definition 2;

$$= \int_{-\infty}^{\infty} g(t-u)\,z(u)\,du.$$

In other words, if G is invariant and linear, it satisfies (1) for complex inputs as well as real ones. Similarly,

$$Gz(t) = z * g(t) = \int_{-\infty}^{\infty} z(t-u)\,g(u)\,du \tag{2}$$

is valid for complex inputs, $z(t)$.

We may summarize this tedious discussion by saying that the theory of linear systems "looks the same" and "works the same" for real and complex inputs, so we need not, in theoretical work, pay much attention to the distinction.

Transfer Functions and Laplace Transforms

We are finally ready to calculate the response of an invariant linear system, G, to a complex sinusoidal input, $z(t) = e^{st}$. By (2),

$$\begin{aligned} Gz(t) &= \int_{-\infty}^{\infty} z(t-u)\, g(u)\, du \\ &= \int_{-\infty}^{\infty} e^{s(t-u)} g(u)\, du \\ &= \int_{-\infty}^{\infty} e^{st} e^{-su} g(u)\, du \\ &= \left[\int_{-\infty}^{\infty} e^{-su} g(u)\, du \right] e^{st}. \end{aligned}$$

Hence

$$Gz(t) = \left[\int_{-\infty}^{\infty} e^{-su} g(u)\, du \right] z(t). \tag{3}$$

The integral on the right defines the (bilateral) *Laplace transform* of g. It is a function of the complex variable s. We denote it $G(s)$ (or, in other contexts, $\mathscr{L}\{g\}(s)$). Thus

$$G(s) = \mathscr{L}\{g\}(s) = \int_{-\infty}^{\infty} e^{-su} g(u)\, du \tag{4}$$

$$= \text{Laplace transform of } g.$$

(If $g(u) = 0$ for $u < 0$, then

$$G(s) = \int_{0}^{\infty} e^{-su} g(u)\, du.) \tag{4'}$$

In terms of this notation, (3) simplifies to

$$Gz(t) = G(s)\, z(t)$$

or

$$Ge^{st} = G(s)\, e^{st}. \tag{5}$$

Thus we have proved the following fundamental theorem.

THEOREM. *If G is an invariant linear system, then complex sinusoidal inputs produce complex sinusoidal outputs of the same complex frequency, s. The entire effect of the system is to multiply the sinusoid by a frequency dependent factor, $G(s)$. This factor is the Laplace transform of the impulse response, g.*

The factor $G(s)$ is also called the *transfer function* of the system G. This terminology emphasizes the relation, (5), of $G(s)$ to the system, whereas the Laplace transform terminology stresses the relation of $G(s)$ to the impulse response. Our interest in $G(s)$ arises from its role in (5), and this equation provides a powerful method of determining $G(s)$. A typical system is defined by an equation relating $y = Gx$ to x. We simply take $x = e^{st}$ and $y = G(s)\,e^{st}$ in this equation and perform a bit of algebra to isolate $G(s)$.

Examples

Consider first the *RCL* circuit, which, for mathematical purposes, is defined by

$$ay'' + by' + y = x.$$

Plugging in $x = e^{st}$ and $y = G(s)\,e^{st}$, and noting that $y' = sy$ and $y'' = s^2y$, we obtain

$$(as^2 + bs + 1)\,G(s)\,e^{st} = e^{st}$$

or

$$G(s) = (as^2 + bs + 1)^{-1} \qquad (RCL \text{ circuit}, a = LC, b = RC). \tag{6}$$

The *RC* circuit (or "*RC* filter" as it is commonly called) is the special case in which $L = 0$, hence $a = 0$ and

$$G(s) = (bs + 1)^{-1} \qquad (RC \text{ circuit}, b = RC). \tag{7}$$

Next we consider the following systems:

amplifier,	$y = ax;$
shift,	$y(t) = x(t - \tau);$
differentiator,	$y = x';$
integrator,	$y' = x.$

Taking $x = e^{st}$ and $y = G(s)\,e^{st}$ in these equations, and isolating $G(s)$ on the left, we obtain

amplifier,	$G(s) = a;$	(8)
shift,	$G(s) = e^{-\tau s};$	(9)
differentiator,	$G(s) = s;$	(10)
integrator,	$G(s) = s^{-1}.$	(11)

All of the examples considered so far, except for the shift, are special cases of the *general integro-differential system*

$$\sum_{k=0}^{n} q_k d^k y/dt^k = \sum_{j=0}^{m} p_j d^j x/dt^j,$$

where $d^0y/dt^0 = y$ and p_j and q_k are real constants. The technique used above easily yields the following expression for the transfer function:

$$G(s) = P(s)/Q(s), \tag{12}$$

where

$$P(s) = \sum_{j=0}^{m} p_j s^j$$

and

$$Q(s) = \sum_{k=0}^{n} q_k s^k$$

are polynomials. A ratio of polynomials is called a *rational function*, so $G(s)$ is called a *rational transfer function.*

Straightforward calculations yield the transfer functions of the simple composite systems $G + H$ and GH. Clearly

$$\begin{aligned}(G + H)\, e^{st} &= Ge^{st} + He^{st}\\ &= G(s)\, e^{st} + H(s)\, e^{st}\\ &= (G(s) + H(s))\, e^{st},\end{aligned}$$

and

$$\begin{aligned}(GH)\, e^{st} &= G(He^{st})\\ &= G(H(s)\, e^{st})\\ &= H(s)\, Ge^{st}\\ &= H(s)\, G(s)\, e^{st}\\ &= G(s)\, H(s)\, e^{st}.\end{aligned}$$

Thus

$$(G + H)(s) = G(s) + H(s) \tag{13}$$

and

$$(GH)(s) = G(s)\, H(s). \tag{14}$$

The transfer function of a sum is the sum of the component transfer functions, and the transfer function of a cascade is the product of the component transfer functions.

Laplace Transforms of Convolutions

Equation (14) has some interesting corollaries. We saw in the last lecture that the impulse response of GH is $g * h$, where $g = G\delta$ and $h = H\delta$. The transfer function $(GH)(s)$ is the Laplace transform of this impulse response,

$$(GH)(s) = \mathscr{L}\{g * h\}(s),$$

just as $G(s) = \mathscr{L}\{g\}(s)$ and $H(s) = \mathscr{L}\{h\}(s)$. Hence (14) is equivalent to

$$\mathscr{L}\{g * h\}(s) = \mathscr{L}\{g\}(s)\,\mathscr{L}\{h\}(s). \tag{15}$$

The Laplace transform of a convolution is the product of the Laplace transforms.

This is valid for arbitrary functions g and h, since any function is the impulse response of an invariant linear system. Replacing h by x and $g * x = Gx$ by y, we see that

$$\mathscr{L}\{y\}(s) = \mathscr{L}\{g\}(s)\,\mathscr{L}\{x\}(s)$$

or

$$Y(s) = G(s)\,X(s). \tag{16}$$

The Laplace transform of the output is the Laplace transform of the input times the transfer function of the system.

The Laplace Transform is Invertible

The transfer function $G(s)$ determines the response of G to complex sinusoids. If G and H are two invariant linear systems with $G(s) = H(s)$ for all s, then $Gx = Hx$ for complex sinusoids $x = e^{st}$. Must $Gx = Hx$ for *all* inputs x, so that $G = H$? In other words,

Does $G(s) = H(s)$ imply $G = H$?

Now $G(s) = \mathscr{L}\{g\}(s)$, $H(s) = \mathscr{L}\{h\}(s)$, and $G = H$ if and only if $g = h$. Thus our question is equivalent to

Does $\mathscr{L}\{g\} = \mathscr{L}\{h\}$ imply $g = h$?

or

Is the correspondence between g and $\mathscr{L}\{g\}$ "invertible" or "one-to-one"?

In the next lecture, we shall see that the answer to this question is "yes." Hence systems are completely characterized by their transfer functions, and functions are completely characterized by their Laplace transforms. The latter statement applies to input and output functions, as well as impulse responses.

Equation (16) shows that the action of G is very simple at the "level" of Laplace transforms. It now emerges that no information is lost by operating at this level, since $Y(s) = \mathscr{L}\{y\}(s)$ completely determines y. Thus the description (16) of G is both simple and complete.

Exercise

Rederive the transfer function, (6), of the RCL circuit by integration. That is, use (4′) in conjunction with Eq. (13) of Lecture 5. Assume $d \neq 0$ and $\operatorname{re} s > \operatorname{re} s_+$.

Lecture 8. Inversion and Interpretation of Laplace Transforms

This lecture has two main objectives: (1) to show how a function, $x(t)$, can be recovered from its Laplace transform, $X(s)$, and (2) to provide a "concrete interpretation" of $X(s)$. We shall obtain a single formula [Eq. (7) below] that accomplishes both objectives.

Our starting point is the following, seemingly unrelated problem: Given a function, $x(t)$, how can we represent this function as a superposition of complex sinusoids? I have in mind a "continuous superposition,"

$$x(t) = c\int_{-\infty}^{\infty} e^{st}F(s)\,d\omega, \tag{1}$$

where $s = \sigma + i\omega$, and c is some constant, to be chosen later. All of the complex frequencies, $s = \sigma + i\omega$, that appear in this formula have the same real part, σ, but a continuum of imaginary parts occur, corresponding to a continuum of different frequencies of circular motion. The factor $cF(s)\,d\omega$ is the weight given to the complex sinusoid e^{st} in the superposition. The constant real part, σ, plays a rather peripheral role in (1). We will have more to say about it later.

It is an amazing fact that most functions can be decomposed into complex sinusoids in this way. We shall not prove this. *Instead we shall assume that it is correct, and then we shall determine what the appropriate weight function must be.* You will see shortly that the result of this exercise will be quite illuminating.

Assume, then, that (1) is given, and let us try to determine a formula that expresses $F(s)$ in terms of $x(t)$. (This amounts to "inverting" (1).) Let $\tilde{s} = \sigma + i\tilde{\omega}$ be a complex number with the same real part as s, and multiply (1) by $e^{-\tilde{s}t}$:

$$\begin{aligned} e^{-\tilde{s}t}x(t) &= c\int_{-\infty}^{\infty} e^{(s-\tilde{s})t}\,F(s)\,d\omega \\ &= c\int_{-\infty}^{\infty} e^{i(\omega-\tilde{\omega})t}F(s)\,d\omega, \end{aligned}$$

since the real parts of s ans $\tilde{s}$ cancel. Now integrate both sides from $t = -T$ to $t = T$:

$$\int_{-T}^{T} e^{-\tilde{s}t}x(t)\,dt = c\int_{-T}^{T}\left[\int_{-\infty}^{\infty} e^{i(\omega-\tilde{\omega})t}F(s)\,d\omega\right]dt$$
$$= c\int_{-\infty}^{\infty}\left[\int_{-T}^{T} e^{i(\omega-\tilde{\omega})t}F(s)\,dt\right]d\omega,$$

where we have interchanged the orders of ω and t integration;

$$= c\int_{-\infty}^{\infty}\left[\int_{-T}^{T} e^{i(\omega-\tilde{\omega})t}\,dt\right]F(s)\,d\omega,$$

since $F(s) = F(\sigma + i\omega)$ doesn't depend on t;

$$= c\int_{-\infty}^{\infty} \frac{e^{i(\omega-\tilde{\omega})t}}{i(\omega-\tilde{\omega})}\Bigg|_{t=-T}^{T} F(s)\,d\omega$$
$$= c\int_{-\infty}^{\infty} \frac{e^{i(\omega-\tilde{\omega})T} - e^{-i(\omega-\tilde{\omega})T}}{i(\omega-\tilde{\omega})} F(s)\,d\omega$$
$$= c\int_{-\infty}^{\infty} \frac{2\sin[(\omega-\tilde{\omega})T]}{\omega-\tilde{\omega}} F(s)\,d\omega$$
$$= 2c\int_{-\infty}^{\infty} \frac{\sin[(\omega-\tilde{\omega})T]}{(\omega-\tilde{\omega})T} F(\sigma+i\omega)T\,d\omega.$$

Now make the change of variables $u = T(\omega - \tilde{\omega})$, and note that $du = T\,d\omega$ and $\omega = \tilde{\omega} + T^{-1}u$, so that

$$\int_{-T}^{T} e^{-\tilde{s}t}x(t)\,dt = 2c\int_{-\infty}^{\infty} \frac{\sin u}{u} F(\tilde{s} + iT^{-1}u)\,du.$$

If we now let $T \to \infty$, we obtain

$$\int_{-\infty}^{\infty} e^{-\tilde{s}t}x(t)\,dt = 2c\int_{-\infty}^{\infty} \frac{\sin u}{u} F(\tilde{s})\,du,$$

since $T^{-1}u \to 0$, or

$$\int_{-\infty}^{\infty} e^{-\tilde{s}t}x(t)\,dt = 2c\left[\int_{-\infty}^{\infty} \frac{\sin u}{u}\,du\right]F(\tilde{s}). \tag{2}$$

Reference to a table of definite integrals shows that

$$\int_{-\infty}^{\infty} \frac{\sin u}{u}\,du = \pi.$$

Substituting this into (2), we see that

$$F(\bar{s}) = \frac{1}{2\pi c} \int_{-\infty}^{\infty} e^{-\bar{s}t} x(t)\, dt,$$

or

$$F(s) = \frac{1}{2\pi c} \int_{-\infty}^{\infty} e^{-st} x(t)\, dt,$$

or

$$F(s) = \frac{1}{2\pi c} \mathscr{L}\{x\}(s). \tag{3}$$

We have not yet decided how to choose the constant factor c. We are at liberty to choose it in any way we please. The choice suggested by (3) is

$$c = 1/2\pi,$$

in which case (1) reduces to

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{st} F(s)\, d\omega \tag{4a}$$

and (3) becomes

$$F(s) = \mathscr{L}\{x\}(s). \tag{4b}$$

To recapitulate this lengthy development: *If a function $x(t)$ can be decomposed, as in* (4a), *into a superposition of complex sinusoids, then the contribution of the complex frequency $s = \sigma + i\omega$ to the superposition is proportional to $\mathscr{L}\{x\}(s)$. This is the promised concrete interpretation of the Laplace transform.*

There remains the awkward "if" in the previous statement but one. It remains to state a condition under which this qualification can be removed, so that we can assert, unconditionally, that (4a) and (4b) are valid. The need for some sort of condition is clear if we consider the formula for the Laplace transform:

$$\mathscr{L}\{x\}(s) = \int_{-\infty}^{\infty} e^{-st} x(t)\, dt. \tag{5}$$

Nothing said so far guarantees that this integral is well defined! We shall certainly have to make an assumption that will ensure that the Laplace transform makes sense. In addition, we shall have to assume that the function x is fairly smooth near the points, t, to which (4a) applies.

THEOREM. *Suppose that, for some constants T, K, and σ_0, we have $x(t) = 0$ for $t < T$ and*

$$|x(t)| \leqslant Ke^{\sigma_0 t} \qquad \textit{for} \quad t \geqslant T. \tag{6}$$

(We say that x is "of exponential order.") Then $X(s) = \mathscr{L}\{x\}(s)$ is well defined if $\sigma = \operatorname{re} s > \sigma_0$. Moreover

$$x(t) = (2\pi)^{-1}\int_{-\infty}^{\infty} e^{st}X(s)\, d\omega, \qquad s = \sigma + i\omega, \tag{7}$$

provided that the derivative, $x'(u)$, exists and is continuous in some small interval, $t - \varepsilon < u < t + \varepsilon$, around t.

Of course, (7) is just a combination of (4a) and (4b). I will skip the proof of (7), since it has much in common with the argument that led to (4a) and (4b). Thus it remains only to show that $\mathscr{L}\{x\}(s)$ is well defined.

Proof that $\mathscr{L}\{x\}(s)$ is well defined. It suffices to show that the integral in (5) is finite when the integrand is replaced by its absolute value, $|e^{-st}|\,|x(t)|$. But

$$e^{-st} = e^{-\sigma t}e^{-i\omega t},$$

so

$$|e^{-st}| = e^{-\sigma t}.$$

Since x is of exponential order,

$$|e^{-st}|\,|x(t)| \leqslant Ke^{-(\sigma-\sigma_0)t} \qquad \text{for} \quad t \geqslant T,$$

and $|e^{-st}|\,|x(t)|$ is zero for $t < T$. Thus

$$\int_{-\infty}^{\infty} |e^{-st}|\,|x(t)|\, dt \leqslant K\int_{T}^{\infty} e^{-(\sigma-\sigma_0)t}\, dt$$

$$= K(\sigma - \sigma_0)^{-1}\, e^{-(\sigma-\sigma_0)T},$$

provided that $\sigma > \sigma_0$. Thus the integral is finite, as was to be shown.

Inverse Laplace Transform

Not only does (7) tell us what the Laplace transform "means," but it also shows that a function, x, can be reconstructed from its Laplace transform, X. In other words, (7) shows that the Laplace transform is invertible—that there is a one-to-one correspondence between functions and their Laplace transforms. To make this clearer, let me define a new transformation, $\mathscr{L}^{-1}$, by

$$\mathscr{L}^{-1}\{X\}(t) = (2\pi)^{-1}\int_{-\infty}^{\infty} e^{st}X(s)\, d\omega, \qquad s = \sigma + i\omega. \tag{8}$$

Then (7) says that

$$x(t) = \mathscr{L}^{-1}\{\mathscr{L}\{x\}\}(t)$$

or

$$x = \mathscr{L}^{-1}\{\mathscr{L}\{x\}\}. \tag{9}$$

The transformation $\mathscr{L}^{-1}$ "undoes" or "inverts" $\mathscr{L}$. We will refer to $\mathscr{L}^{-1}\{X\}$ as the *inverse Laplace transform* of X.

It follows immediately from the definitions of $\mathscr{L}$ and $\mathscr{L}^{-1}$ that both are linear transformations,

$$\mathscr{L}\left\{\sum_{j=1}^{n} a_j x_j\right\} = \sum_{j=1}^{n} a_j \mathscr{L}\{x_j\} \tag{10}$$

and

$$\mathscr{L}^{-1}\left\{\sum_{j=1}^{n} a_j X_j\right\} = \sum_{j=1}^{n} a_j \mathscr{L}^{-1}\{X_j\}. \tag{11}$$

These equations will be used frequently in subsequent lectures.

Inversion without Integration

The integral definition, (8), of $\mathscr{L}^{-1}$ is somewhat forbidding, and we shall, in fact, have little use for it in the future. Laplace transforms are more easily inverted by "reading the Laplace transform table backward."

EXAMPLE 1: THE INTEGRATOR. Let G be the integrator, defined by $y' = x$. In the exercise for Lecture 3, you showed that g is the unit step function

$$\begin{aligned} g(t) = \mu(t) &= 1 \qquad \text{for} \quad t \geqslant 0, \\ &= 0 \qquad \text{for} \quad t < 0. \end{aligned}$$

Moreover, we saw in the last lecture that

$$\mathscr{L}\{g\}(s) = G(s) = s^{-1}.$$

Therefore,

$$\mathscr{L}\{\mu\}(s) = s^{-1}. \tag{12}$$

Applying $\mathscr{L}^{-1}$ to both sides and using (9), we find that $\mu = \mathscr{L}^{-1}\{s^{-1}\}$ or

$$\mathscr{L}^{-1}\{s^{-1}\} = \mu. \tag{13}$$

Note that our derivation of this formula circumvents the integration in (8).

EXAMPLE 2: *RC* FILTER. If G is an RC filter, then

$$\begin{aligned} g(t) &= b^{-1}e^{-t/b}, && \text{for} \quad t \geqslant 0, \\ &= 0, && \text{for} \quad t < 0, \end{aligned}$$

or, more compactly,

$$g(t) = b^{-1}e^{-t/b}\mu(t).$$

Also, we saw in the last lecture that

$$\mathscr{L}\{g\}(s) = G(s) = (bs+1)^{-1},$$

so

$$\mathscr{L}\{b^{-1}e^{-t/b}\mu(t)\}(s) = (bs+1)^{-1}. \tag{14}$$

Just as in Example 1, we conclude that

$$\mathscr{L}^{-1}\{(bs+1)^{-1}\}(t) = b^{-1}e^{-t/b}\mu(t). \tag{15}$$

EXAMPLE 3: CASCADE OF *RC* FILTERS. Let G be a cascade of n RC filters with the same time constant, b, and let g_n be its impulse response. Then

$$\mathscr{L}\{g_n\}(s) = G(s) = (bs+1)^{-n}, \tag{16}$$

since cascading systems corresponds to multiplying transfer functions. Hence

$$\mathscr{L}^{-1}\{(bs+1)^{-n}\}(t) = g_n(t). \tag{17}$$

We shall show next time that

$$g_n(t) = b^{-n}\frac{t^{n-1}}{(n-1)!}e^{-t/b}\mu(t), \tag{18}$$

where $(n-1)!$ is the product of the first $n-1$ positive integers.

These examples illustrate how formulas for Laplace transforms yield equivalent formulas for inverse Laplace transforms.

Interpretation of $Y(s) = G(s)X(s)$

Suppose now that G is an invariant linear system, and let $y = Gx$. Last time we derived the rather mysterious formula

$$Y(s) = G(s)X(s). \tag{19}$$

What does this formula mean, and where does it come from? It says that the amount of e^{st} in y is $G(s)$ times the amount of e^{st} in x. Why is this? According to (7), x is a superposition of complex sinusoids with weighting factors $X(s)$. If we apply G to both sides of (7) and use the linearity of G, we see that

$$y(t) = (2\pi)^{-1} \int_{-\infty}^{\infty} [Ge^{st}]\, X(s)\, d\omega$$

($X(s)$ doesn't depend on t, so it is just a constant as far as G is concerned). But

$$Ge^{st} = e^{st}G(s),$$

so

$$y(t) = (2\pi)^{-1} \int_{-\infty}^{\infty} e^{st}[G(s)\, X(s)]\, d\omega.$$

Equation (19) is obtained by comparing this with

$$y(t) = (2\pi)^{-1} \int_{-\infty}^{\infty} e^{st} Y(s)\, d\omega.$$

This derivation shows very concretely that, if we know how an invariant linear system transforms complex sinusoids, then we know how it transforms anything else. This is a most endearing property of invariant linear systems.

Lecture 9. Rational Transfer Functions, Feedback

Almost all of the systems considered so far have rational transfer functions,

$$G(s) = P(s)/Q(s), \tag{1}$$

where $P(s)$ and $Q(s)$ are polynomials. The importance of such transfer functions is greatly enhanced by the principle that *interconnection of components with rational transfer functions will produce a network whose overall transfer function is rational.*

EXAMPLE: MANUAL TRACKING WITH FEEDBACK. One of the great strengths of linear systems theory is the ease with which it handles systems involving feedback. Consider, for example, the situation depicted in Fig. 19. As in Fig. 16, y is the displacement of the mass from its resting position. However, x is now the vertical position of a moving pointer positioned by an experimenter. A human subject can observe y and x, and can push or pull the mass via an attached rod. He is instructed to keep the pointers as close as he can. This is an example of a *manual tracking experiment.* See Licklider (1960) for a full discussion of manual tracking experiments and models.

Our simple-minded model for the subject's behavior assumes that the force, f, that he exerts on the mass is proportional to the tracking error, $x - y$. Thus $f = \gamma(x - y)$,

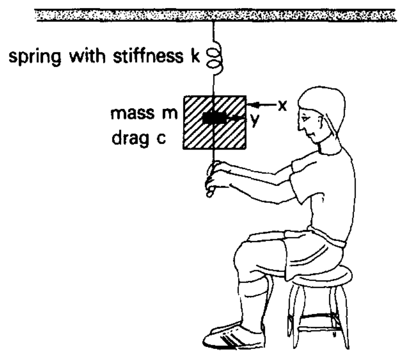

FIG. 19. Manual tracking.

where $\gamma > 0$. The subject pushes the block toward the pointer, and pushes harder when the error is greater. As was noted in Lecture 5, the mechanical system, call it H, that maps force into displacement ($y = Hf$) is described by the differential equation

$$my'' + cy' + ky = f. \tag{2}$$

The man–machine system, G, that maps x into y can be diagrammed as in Fig. 20. The output, y, is "fed back" and subtracted from x. The input of H is proportional to this difference. The equation representing this state of affairs is

$$y = H\gamma(x - y)$$

or

$$y = \gamma H(x - y). \tag{3}$$

We wish to express $G(s)$ in terms of $H(s)$. This is surprisingly easy to do. Taking Laplace transforms on both sides of (3), we obtain

$$Y(s) = \gamma H(s)(X(s) - Y(s)).$$

Thus

$$Y(s)(1 + \gamma H(s)) = \gamma H(s)\, X(s)$$

or

$$Y(s) = \frac{\gamma H(s)}{1 + \gamma H(s)} X(s). \tag{4}$$

Comparing this with $Y(s) = G(s)\, X(s)$, we see that

$$G(s) = \frac{\gamma H(s)}{1 + \gamma H(s)}. \tag{5}$$

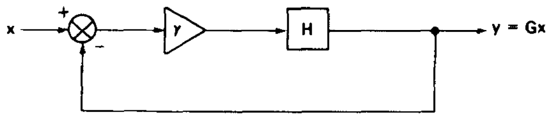

FIG. 20. A negative feedback model for manual tracking.

Note that we could have saved a step by simply treating x, y, and H as numbers in (3), and solving for y in terms of x. This yields

$$y = \frac{\gamma H}{1 + \gamma H} x,$$

which is analogous to (4). The ratio on the right is $G(s)$, shorn of its s's. The notation in this theory is designed to permit such short cuts.

The technique described in Lecture 7, applied to (2), shows immediately that $H(s)$ is the rational function

$$H(s) = (ms^2 + cs + k)^{-1}. \tag{6}$$

Plugging this into (5) and multiplying numerator and denominator by the polynomial in the denominator of $H(s)$, we see that $G(s)$ is the rational function

$$G(s) = \gamma(ms^2 + cs + k + \gamma)^{-1}. \tag{7}$$

This example illustrates the principle that a network will have a rational transfer function whenever its components have rational transfer functions. A more complicated feedback system illustrating this principle is given in the first exercise.

Determining $g(t)$ when $G(s)$ Is Rational

It should now be clear that rational transfer functions are of sufficient importance to justify considerable subsequent attention in these lectures. In the remainder of this lecture, we shall see how to determine the impulse response, g, for a nonanticipating system, G, with transfer function $G(s) = P(s)/Q(s)$. We shall always assume that $P(s)$ is of lower degree that $Q(s)$. This corresponds to the condition that $g = G\delta$ is a bona fide function. We shall see that the roots of Q figure prominently in our formulas for g. Thus these roots must be known before $g(t)$ can be calculated.

Suppose that

$$Q(s) = \sum_{k=0}^{n} q_k s^k \tag{8}$$

has degree $n \geqslant 1$. Let $r_1, \ldots, r_\nu$ be the (distinct) roots of $Q(s)$. Then

$$Q(s) = q_n \prod_{j=1}^{\nu} (s - r_j)^{n_j}, \tag{9}$$

where n_j is a positive integer called the *multiplicity* of r_j, and

$$\sum_{j=1}^{v} n_j = n.$$

If $n_j = 1$, then r_j is called a *simple root*; if $n_j = 2$, it is a *double root*, etc. I will consider successively three cases,

Case 1: n simple roots,

Case 2: cascade of n identical RC filters,

Case 3: general case.

Most applications fall within Case 1, and the formula for $g(t)$ in that case is both simple and easy to derive. In Case 2, $Q(s)$ has a single root of multiplicity n. This system is of considerable interest, and, in addition, it provides a basic formula needed in Case 3. As you might expect, the formula for $g(t)$ in the general case is rather complicated.

Case 1: *n simple roots.* In this case,

$$Q(s) = q_n \prod_{j=1}^{n} (s - r_j), \tag{10}$$

and it can be shown that $G(s)$ has a *partial fractions expansion* of the form

$$G(s) = \sum_{j=1}^{n} a_j/(s - r_j). \tag{11}$$

Granted this, it is easy to see what the coefficients, a_j, must be. Multiplying (11) by $s - r_j$, we find that

$$(s - r_j)\, G(s) = a_j + \sum_{k \neq j} a_k (s - r_j)/(s - r_k).$$

(Note the change of summation index from j to k.) But

$$(s - r_j)/(s - r_k) \to 0 \qquad \text{as} \quad s \to r_j,$$

since $r_j \neq r_k$. Thus

$$a_j = \lim_{s \to r_j} (s - r_j)\, G(s). \tag{12a}$$

However, in view of (10),

$$(s - r_j)\, G(s) = \frac{P(s)}{q_n \prod_{k \neq j} (s - r_k)}.$$

Thus (12a) is equivalent to

$$a_j = \frac{P(r_j)}{q_n \prod_{k \neq j} (r_j - r_k)}. \tag{12b}$$

Returning to (11), taking the inverse Laplace transform of both sides, and recalling that $g = \mathscr{L}^{-1}\{G\}$ and that $\mathscr{L}^{-1}$ is linear, we obtain the expression

$$g(t) = \sum_{j=1}^{n} a_j \mathscr{L}^{-1}\{(s - r_j)^{-1}\}. \tag{13}$$

But, in the course of working the exercise for Lecture 7, you will have shown that

$$\mathscr{L}\{\mu(t)\, e^{rt}\}(s) = (s - r)^{-1}, \tag{14}$$

provided that re $s >$ re r. Thus

$$\mathscr{L}^{-1}\{(s - r)^{-1}\} = \mu(t)\, e^{rt}. \tag{15}$$

This result, in conjunction with (13), yields

$$g(t) = \mu(t) \sum_{j=1}^{n} a_j \exp(r_j t), \tag{16}$$

where $\exp(z) = e^z$. Thus we have proved the following theorem.

THEOREM 1. *Equation* (16) *gives the impulse response corresponding to* $G(s) = P(s)/Q(s)$, *when the roots*, $r_1, r_2, \ldots, r_n$, *of* $Q(s)$ *are all simple. The coefficients* a_j *are given by the alternative formulas* (12a) *and* (12b).

Case 2: *Cascade of n identical RC filters.* The system to be considered can be diagrammed as in Fig. 21. Each stage is an *RC* filter with transfer function $(bs + 1)^{-1}$. The cascade, G_n, is called an *n*-stage *RC filter.* Clearly

$$G_n(s) = (bs + 1)^{-n}. \tag{17}$$

The denominator has one root, $r = -b^{-1}$, of multiplicity n.

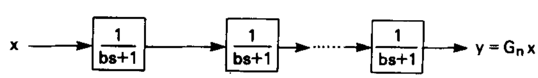

FIG. 21. An *n*-stage *RC* filter.

THEOREM 2. *The impulse response, $g_n = G_n\delta$, is*

$$g_n(t) = \mu(t)\, b^{-n} \frac{t^{n-1}}{(n-1)!} e^{-t/b}. \tag{18}$$

For $n > 1$, $g_n(0) = 0$, $g_n(t)$ rises to a peak at $t_p = (n-1)b$, and then falls off to 0. The case $n = 3$ is illustrated in Fig. 22. To see that $g_n(t) \to 0$ as $t \to \infty$, recall that

$$\begin{aligned} e^{t/b} &= \sum_{j=0}^{\infty} (t/b)^j/j! \\ &\geqslant (t/b)^n/n!, \end{aligned}$$

hence

$$e^{-t/b} \leqslant n!\, b^n/t^n$$

and

$$g_n(t) \leqslant n/t$$

for $t > 0$. Clearly $n/t \to 0$ as $t \to \infty$, so $g_n(t) \to 0$ too.

Proof of Theorem 2 *by mathematical induction.* When $n = 1$, (18) reduces to

$$g_1(t) = \mu(t)\, b^{-1} e^{-t/b},$$

which is correct, since the expression on the right is, indeed, the impulse response of a single RC filter. Thus (18) is valid for $n = 1$.

My next move may surprise you. *I shall show that validity of* (18) *for some*

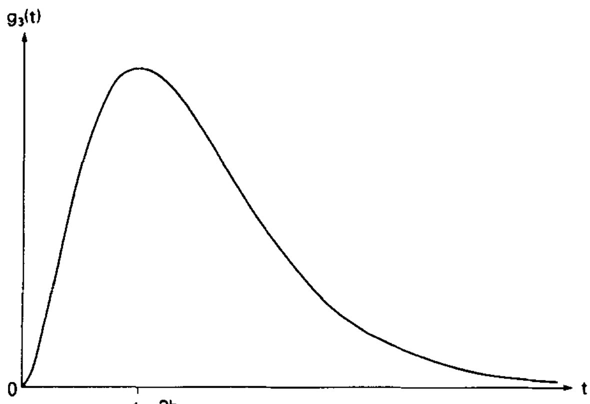

FIG. 22. Impulse response of a 3-stage RC filter.

integer, say, $n=8$, *implies its validity for the next integer,* $n=9$. A cascade of 9 systems is just a cascade of 8 cascaded with one more. Moreover, cascades of systems correspond to convolutions of impulse responses. Thus $g_9 = g_1 * g_8$. Since we are assuming that (18) holds for $n=8$ and we know that it holds for $n=1$, it follows that

$$\begin{aligned} g_9(t) &= \int_{-\infty}^{\infty} g_1(t-u)\, g_8(u)\, du \\ &= \frac{b^{-9}}{7!} \int_0^t e^{-(t-u)/b} u^7 e^{-u/b}\, du \\ &= \frac{b^{-9}}{7!} e^{-t/b} \int_0^t u^7\, du \\ &= \frac{b^{-9}}{8!} t^8 e^{-t/b} \end{aligned}$$

for $t \geqslant 0$. Hence validity of (18) for $n=8$ implies its validity fo $n=9$, as claimed. There is, of course, nothing special about 8 and 9.

The remainder of the proof is like the fall of a row of dominos. We know that (18) is valid for $n=1$. Thus it is valid for $n=2$. Thus it is valid for $n=3$, thus for $n=4$, and so on for all positive integers. This type of argument is called *mathematical induction.* This completes the proof of Theorem 2.

We may combine (17) and (18) into a single equation as follows,

$$\mathscr{L}\left\{\mu(t)\, b^{-n} \frac{t^{n-1}}{(n-1)!} e^{-t/b}\right\}(s) = (bs+1)^{-n}. \tag{19}$$

Multiplying both sides by b^n and recalling that $\mathscr{L}$ is linear, we obtain

$$\mathscr{L}\left\{\mu(t) \frac{t^{n-1}}{(n-1)!} e^{-t/b}\right\}(s) = (s+b^{-1})^{-n}.$$

If we let $r = -b^{-1}$, this becomes

$$\mathscr{L}\left\{\mu(t) \frac{t^{n-1}}{(n-1)!} e^{rt}\right\}(s) = (s-r)^{-n}. \tag{20}$$

Our derivation establishes (20) only for r real and negative; however, it can be shown by other methods that (20) is valid for arbitrary complex numbers, r, provided that $\operatorname{re} s > \operatorname{re} r$. Equation (20) is equivalent to

$$\mathscr{L}^{-1}\{(s-r)^{-n}\} = \mu(t) \frac{t^{n-1}}{(n-1)!} e^{rt}. \tag{21}$$

This equation is the key to our handling of the general rational transfer function.

Case 3: *General case.* Here we place no restriction on the multiplicities, n_j, of the roots, r_j, of $Q(s)$. In this case, $G(s) = P(s)/Q(s)$ admits a partial fractions expansion

$$G(s) = \sum_{j=1}^{\nu} \sum_{k=1}^{n_j} a_{jk}(s - r_j)^{-k},$$

which generalizes (11). Applying $\mathscr{L}^{-1}$ to both sides, and using the linearity of $\mathscr{L}^{-1}$ and Eq. (21), we obtain

$$g(t) = \sum_{j=1}^{\nu} \sum_{k=1}^{n_j} a_{jk}\mu(t) \frac{t^{k-1}}{(k-1)!} \exp(r_j t), \tag{22}$$

or

$$g(t) = \mu(t) \sum_{j=1}^{\nu} P_j(t) \exp(r_j t), \tag{23}$$

where

$$P_j(t) = \sum_{k=1}^{n_j} a_{jk} \frac{t^{k-1}}{(k-1)!} \tag{24}$$

is a polynomial. We will give a formula for a_{jk} at the end of the lecture. The detailed expression, (24), for $P_j(t)$ will not be needed in our subsequent work. However, we shall make essential use of (22) and (23). The following theorem summarizes our results.

THEOREM 3. *If $G(s) = P(s)/Q(s)$, where $P(s)$ is of lower degree than $Q(s)$, then $g(t)$ is given by* (23), *where $P_j(t)$ is a polynomial and $r_1, \ldots, r_\nu$ are the roots of $Q(s)$.*

A difficulty with (23) and even with the special case (16) is that these formulas presuppose that the roots of $Q(s)$ are known precisely, whereas, if the degree of $Q(s)$ exceeds two, it may be difficult or impossible to determine these roots. As an indication of this problem, I mention that the analogs of the quadratic formula for cubic and quartic polynomials are much more complicated, and there is no such formula for polynomials of degree 5 or greater. In the next lecture we shall see that certain important qualitative properties of $g(t)$ can be inferred directly from $G(s)$ without first deriving an explicit formula for $g(t)$.

Two Technicalities

1. *Formulas for a_{jk} in Case* 3. Let

$$G_j(s) = (s - r_j)^{n_j} G(s).$$

It can be shown that

$$a_{jn_j} = G_j(r_j),$$

and that, for $1 \leqslant k < n_j$,

$$a_{jk} = (l!)^{-1}\,(d^l/ds^l)\,G_j(r_j),$$

where $l = n_j - k$. These formulas are stated for the sake of completeness. We will not need them in these lectures.

2. *Phony roots.* Suppose that

$$P(s) = s + 2$$

and

$$Q(s) = s^2 + 3s + 2 = (s+1)(s+2).$$

The roots of $Q(s)$ are $r_1 = -1$ and $r_2 = -2$, and Theorem 1 yields

$$g(t) = \mu(t)(a_1 e^{-t} + a_2 e^{-2t}), \tag{25}$$

where $a_1 = 1$ and $a_2 = 0$; that is,

$$g(t) = \mu(t)\, e^{-t}. \tag{26}$$

Obviously (25) conveys a misleading impression of the nature of $g(t)$, whereas (26) clears the matter up. We could have arrived at (26) more directly if we had cancelled the "phony factor," $s + 2$, of $P(s)$ and $Q(s)$, corresponding to the "phony root," -2, of $Q(s)$.

Generalizing from this example, suppose that $P(s)$ and $Q(s)$ have no common roots. Then, in Case 1, all of the coefficients, a_j, in (16) are nonzero. In Case 3, all of the leading coefficients, a_{jn_j}, of the polynomials $P_j(t)$ are nonzero, so the degree of $P_j(t)$ is $n_j - 1$.

Exercises

1. Consider the feedback system, G, defined by Fig. 23. Express $G(s)$ in terms of $H_1(s)$ and $H_2(s)$. Show that $G(s)$ is rational if both $H_1(s)$ and $H_2(s)$ are rational.

2. Find $h(t)$ and $g(t)$ corresponding to the transfer functions given in (6) and (7). Can $g(t)$ have an oscillatory factor ($\sin \omega t$) even if $h(t)$ doesn't? Must $g(t)$ have such a factor if $h(t)$ does?

3. Let $y(t)$ be the voltage across the capacitor in an RCL circuit, let $q(t)$ be the charge on the capacitor, and let $i(t)$ be the current in the circuit. We know that $y(t) = q(t)/C$, hence $y' = q'/C$, or $i = Cy'$. We can consider $i(t)$ as the output of the RCL circuit corresponding to the voltage input, $x(t)$. This defines an invariant linear system M, $i = Mx$. Determine $M(s)$ and $m(t)$. (Hint: M is a cascade of the systems $y = Gx$ and $i = Cy'$.)

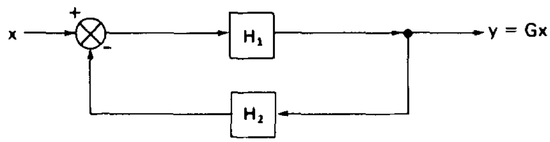

FIG. 23. A feedback system.

Lecture 10. *Determining the Shape of g*

The formulas for $g = G\delta$ developed in the last lecture may be difficult or impossible to apply in certain cases. Moreover, even when we can apply these formulas, the resulting expressions may not immediately suggest the shape of g. Thus it is highly desirable to find ways of inferring the shape of g directly from $G(s)$, without actually deriving an explicit formula for g. That will be our main goal in this lecture. As in Lecture 9, we assume that G is nonanticipating.

One topic to be considered is stability. We say that G (or g) is *stable* if $g(t) \to 0$ as $t \to \infty$. Otherwise it is *unstable*. Clearly stability or instability is an important component of the shape of g. Another component to be considered is the amount of oscillation of g, indexed by the number of zeros of g. Finally, to compose a picture of g, we must know whether or not $g(0) = 0$. We shall begin by obtaining two simple formulas for $g(0)$.

Formulas for $g(0)$

Suppose that g is of exponential order, so that its Laplace transform,

$$G(s) = \mathscr{L}\{g\}(s) = \int_0^\infty e^{-st} g(t)\, dt, \tag{1}$$

is well defined for sufficiently large real values of s. (Review the theorem in Lecture 8, if necessary.)

THEOREM 1. $g(0) = \lim_{s\to\infty} sG(s)$.

Proof. Clearly

$$\begin{aligned} sG(s) &= \int_0^\infty s e^{-st} g(t)\, dt \\ &= \int_0^\infty b^{-1} e^{-t/b} g(t)\, dt, \end{aligned}$$

where I have written b^{-1} in place of s. The term involving b is, of course, the impulse response, $g_b(t)$, of an RC circuit. The area under this function is unity and the function decreases by a factor e^{-1} as t goes from 0 to its time constant, b. See Fig. 24. As $b \to 0$ (or $s = b^{-1} \to \infty$), g_b approaches δ, so

$$\begin{aligned} sG(s) &= \int_0^\infty g_b(t)\, g(t)\, dt \\ &\to \int_0^\infty \delta(t)\, g(t)\, dt \\ &= g(0). \end{aligned}$$

Thus $sG(s) \to g(0)$, as was to be shown.

Throughout the remainder of the lecture, we will assume that $G(s) = P(s)/Q(s)$ is rational and (as usual) $n > m$, where m and n are the degrees of

$$P(s) = \sum_{j=0}^{m} p_j s^j \tag{2}$$

and

$$Q(s) = \sum_{k=0}^{n} q_k s^k. \tag{3}$$

COROLLARY OF THEOREM 1. *If $n = m + 1$, then $g(0) = p_m/q_n$. If $n > m + 1$, then $g(0) = 0$.*

Proof. For large s, the leading terms, $p_m s^m$ and $q_n s^n$, of $P(s)$ and $Q(s)$ predominate, so

$$G(s) \sim (p_m s^m)/(q_n s^n)$$

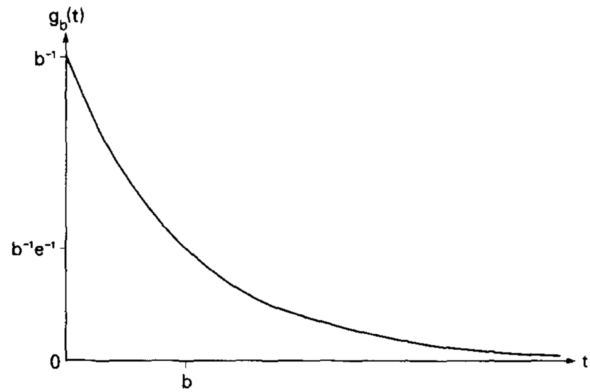

FIG. 24. Impulse response of an RC circuit.

and

$$sG(s) \sim (p_m/q_n)/s^{n-m-1}$$

as $s \to \infty$. Thus

$$g(0) = \lim_{s \to \infty} sG(s) = p_m/q_n$$

if $n = m + 1$, and $g(0) = 0$ if $n > m + 1$, as was to be shown.

Stability

Recall that stability means $g(t) \to 0$ as $t \to \infty$. The system eventually quiets down after it is "kicked." It tuns out that G is stable if and only if all the roots of Q have negative real parts. This is very surprising the first time one hears it, but the significance of negative real parts becomes immediately apparent in the course of the proof.

THEOREM 2. *G is stable if and only if the roots of $Q(s)$ all have negative real parts.*

I shall only prove that negative real parts imply stability. The converse statement, stability implies negative real parts, does not, by the way, apply to phony roots. If $g(t)$ is stable, one can infer only that the "essential" roots of $Q(s)$ (those not common with $P(s)$) have negative real parts.

Proof that negative real parts imply stability. Equation (22) of the last lecture yields

$$g(t) = \sum_{j=1}^{\nu} \sum_{k=1}^{n_j} a_{jk} \frac{t^{k-1}}{(k-1)!} \exp(r_j t) \tag{4}$$

for $t > 0$. Moreover,

$$|\exp(r_j t)| = \exp(\sigma_j t),$$

where $\sigma_j = \text{re } r_j$. Since $\sigma_j < 0$,

$$t^{k-1} \exp(\sigma_j t) \to 0$$

as $t \to \infty$ (see the discussion following Theorem 2 of the last lecture) for all j and k. Hence $g(t) \to 0$, as was to be shown.

I have already mentioned the difficulty of finding roots of some polynomials of degree greater than 2. Fortunately, it is relatively easy to determine whether the roots all have negative real parts, using an algorithm due to Routh (see Schwarz & Friedland, 1965, pp. 406–408). This algorithm can be applied to polynomials of any degree. For polynomials of degree 2, 3, and 4 it leads to the following theorem.

THEOREM 3. *Suppose that $Q(s)$ is given by* (3), *where $n = 2$, 3, or 4. Then all roots of $Q(s)$ have negative real parts if and only if the following conditions are satisfied.*

$n = 2$: *All q_k's have the same sign.*

$n = 3$: *All q_k's have the same sign, and*

$$q_1 q_2 > q_0 q_3. \tag{5}$$

$n = 4$: *All q_k's have the same sign, and*

$$q_2 q_3 - q_1 q_4 > q_3^2 q_0 / q_1. \tag{6}$$

For a more thorough discussion of stability, see Chapters 11 and 12 of Schwarz and Friedland (1965).

Stability and Feedback

Some interesting applications of this theory arise in the context of the simple feedback scheme shown in Fig. 20. Our derivation showed that, for any invariant linear system H,

$$G(s) = \frac{\gamma H(s)}{1 + \gamma H(s)}. \tag{7}$$

If H is a simple mechanical system, as in a manual tracking model,

$$H(s) = (ms^2 + cs + k)^{-1}, \tag{8}$$

so

$$G(s) = \gamma(ms^2 + cs + k + \gamma)^{-1}. \tag{9}$$

Assuming, as usual, that m, c, k, and γ are positive, all roots of the denominators of $H(s)$ and $G(s)$ have negative real parts, according to Theorem 3. Thus, by Theorem 2, both H and G are stable.

The feedback in this example is negative. A comparable system with positive feedback is obtained by changing the minus to a plus on the left in Fig. 20. (Obviously the new system is not a model for manual tracking!) Proceeding just as in the case of negative feedback, we obtain

$$G(s) = \gamma(ms^2 + cs + k - \gamma)^{-1} \tag{10}$$

for the mechanical system with positive feedback. The coefficients in the denominator have the same sign if and only if $k > \gamma$; hence, by Theorems 2 and 3, the positive feedback system is stable if and only if the amplification factor, γ, is less than the stiffness, k, of the spring.

Returning to the system with negative feedback, one is tempted to speculate that stability of H implies stability of G. This generalization is incorrect. Suppose, for example, that H is a three-stage RC filter, so that

$$H(s) = (bs + 1)^{-3}. \tag{11}$$

The single root, $-b^{-1}$, of the denominator is negative, so H is stable. Plugging (11) into (7) and simplifying, we obtain

$$\begin{aligned} G(s) &= \gamma[(bs + 1)^3 + \gamma]^{-1} \\ &= \gamma[b^3s^3 + 3b^2s^2 + 3bs + 1 + \gamma]^{-1}. \end{aligned}$$

All of the coefficients in the denominator are positive, so, by Theorem 2 and 3, G is stable if and only if (5) holds, that is,

$$(3b)(3b^2) > (1 + \gamma)b^3.$$

Cancelling b^3, we see that G is stable if $\gamma < 8$. If $\gamma \geqslant 8$, then G is unstable, even though feedback is negative and H is stable.

A Formula for $\int_0^\infty g(t)\,dt$

For stable systems with rational transfer functions, it can be shown that $g(t)$ converges to 0 (as $t \to \infty$) sufficiently rapidly that

$$\int_0^\infty |g(t)|\,dt < \infty. \tag{12}$$

If g is any impulse response satisfying (12), the integral in (1) is well defined for all s with nonnegative real part. Taking $s = 0$, we obtain the useful formula

$$\int_0^\infty g(t)\,dt = G(0) \tag{13}$$

for the integral of the impulse response. Recalling that

$$G\mu(t) = \int_0^t g(u)\,du \tag{14}$$

(see (9) in Lecture 6), we see that (13) is equivalent to the formula

$$G\mu(\infty) = G(0) \tag{15}$$

for the asymptote of the step response, $G\mu$.

Oscillation

Stability or instability is an important aspect of the shape of g. Another is the degree to which g oscillates or "rings." A straightforward measure of the oscillation of g is the number of times it hits the nonnegative t axis, that is, the number of distinct nonnegative zeros of g. For example, the impulse response shown in Fig. 25 has two nonnegative zeros. Let

$$\#g = \text{number of nonnegative zeros.}$$

We know that $\#g = 0$ for RC circuits. For RCL circuits,

$$\begin{aligned} \#g &= 1, && \text{if} \quad d = b^2 - 4a \geqslant 0, \\ &= \infty, && \text{if} \quad d < 0. \end{aligned}$$

In this example, $\#g < \infty$ when the roots of the characteristic polynomial, $Q(s)$, are real. The next theorem implies that $\#g < \infty$ in any system with rational transfer function for which all roots of $Q(s)$ are real.

THEOREM 4. (Oscillation Theorem). *Suppose that* $G(s) = P(s)/Q(s)$, *where all roots of* $Q(s)$ *are real. Let* $n = \text{degree } Q$. *Then*

$$\#g < n. \tag{16}$$

Furthermore, the same inequality applies to any derivative of g,

$$\begin{aligned} \#g' &< n, \\ \#g'' &< n, \end{aligned} \tag{17}$$

etc.

Clearly, g doesn't oscillate much at all if $Q(s)$ has real roots and low degree.

There is one obvious exception to the theorem. If $P(s) = 0$ for *all* s, so that $G(s) = 0$ for *all* s, then $g(t) = 0$ for *all* t. Thus, to be completely precise, we must explicitly assume that $P(s)$ is not identically zero.

The proof of Theorem 4 is a bit too difficult for inclusion here. It is given in the Appendix at the end of these notes.

Hill's Model

To illustrate the methods introduced in this lecture, we shall analyze a system, G, described by Fig. 26. Here the output, v, of $\gamma_1 H_1$ has a direct path to the final output,

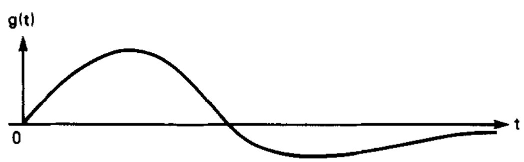

FIG. 25. An impulse response with two nonnegative zeros.

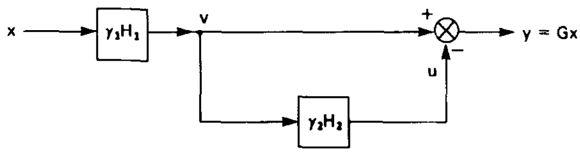

FIG. 26. A model with an inhibitory feedforward path.

y, and an indirect, inhibitory, "feedforward" path through $\gamma_2 H_2$. We shall assume that H_1 and H_2 are RC filters with time constants $b_1 < b_2$. Naturally, γ_1 and γ_2 are positive constants (amplification factors).

A scheme of this sort is Hill's (1936) model for the generation of nerve impulses. In that model, x is stimulating current, v is the (hypothetical) excitation produced thereby, and $u = \gamma_2 H_2 v$ is a (hypothetical) accommodating threshold. More precisely, u and v are departures of the threshold and excitation from resting levels, u_0 and v_0. A nerve impulse is generated when $v + v_0$ hits $u + u_0$, i.e., $y = v - u$ hits $u_0 - v_0$. In addition to $b_1 < b_2$, Hill assumed $\gamma_2 = 1$.

Solomon and Corbit (1974; see Panel B of Fig. 3, p. 126) incorporated a scheme of this kind in their theory of motivation. In their theory, $x(t)$ is the amount of stimulation at time t (e.g., by opiate injection), u and v are opponent motivational processes with opposite affective tone, and $y = v - u$ represents net affect. Solomon and Corbit did not make precise assumptions about H_1 and H_2, and some of their ideas would seem to require that H_2, at least, be nonlinear.

Our analysis of G begins with the calculation of its transfer function. Clearly

$$\begin{aligned} y &= v - u \\ &= \gamma_1 H_1 x - \gamma_1 \gamma_2 H_1 H_2 x, \end{aligned}$$

hence

$$G(s) = \gamma_1 H_1(s)(1 - \gamma_2 H_2(s)),$$

or

$$G(s) = \frac{\gamma_1(b_2 s + 1 - \gamma_2)}{(b_1 s + 1)(b_2 s + 1)}. \tag{18}$$

We shall now see how much information about g can be extracted from (18) via the methods of this lecture. The degrees of the numerator and denominator of $G(s)$ are $m = 1$ and $n = 2$, and the leading coefficients are $p_1 = \gamma_1 b_2$ and $q_2 = b_1 b_2$. Hence the Corollary of Theorem 1 yields

$$g(0) = \gamma_1 / b_1.$$

In particular, $g(0) > 0$.

The roots of the denominator are real and negative, so $g(t) \to 0$ as $t \to \infty$ by Theorem 2, and

$$\#g < 2$$

and

$$\#g' < 2,$$

by Theorem 4. The only four shapes consistent with this information are shown in Fig. 27.

Can some of these shapes be excluded? By (13),

$$\int_0^\infty g(t)\,dt = \gamma_1(1-\gamma_2).$$

If $\gamma_2 \geqslant 1$, the integral is less than or equal to zero, and the first three shapes can be ruled out. In fact, the same is true even if $\gamma_2 < 1$, but to show this we must appeal to Theorem 1 of the last lecture. According to that theorem,

$$g(t) = a_1 \exp(-t/b_1) + a_2 \exp(-t/b_2) \tag{19}$$

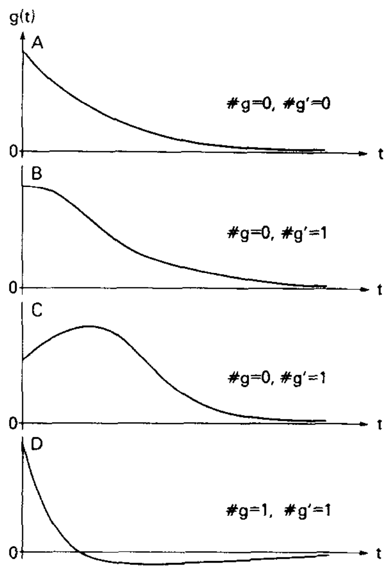

FIG. 27. The four shapes with $g(0) > 0$, $\#g < 2$, $\#g' < 2$, and $g(\infty) = 0$.

for $t \geqslant 0$, where

$$\begin{aligned} a_2 &= \frac{P(-b_2^{-1})}{b_2 - b_1} \\ &= -\frac{\gamma_1 \gamma_2}{b_2 - b_1} \\ &< 0, \end{aligned} \tag{20}$$

since we have assumed $b_2 > b_1$. It follows from (19) and $b_2^{-1} < b_1^{-1}$ that

$$g(t)/\exp(-t/b_2) \to a_2$$

as $t \to \infty$, hence $g(t) < 0$ for t sufficiently large. Consequently g must look like the bottom panel of Fig. 27.

Exercises

All of these problems relate to the Hill–Solomon–Corbit scheme considered above.

1. Use (15) to obtain a formula for $G\mu(\infty)$.
2. According to (14), $G\mu(0) = 0$ and $(G\mu)'(t) = g(t)$. Use these facts, together with the above picture of $g(t)$ and your formula for $G\mu(\infty)$, to sketch $G\mu(t)$ for $\gamma_2 < 1$, for $\gamma_2 = 1$, and for $\gamma_2 > 1$.
3. Work out a formula for a_1 analogous to (20).
4. Use (14) and (19) to obtain a formula for $G\mu(t)$.
5. Take $\gamma_1 = 1$, $\gamma_2 = 0.5$, $b_1 = 0.1$, and $b_2 = 0.2$ in your formula for $G\mu(t)$, and tabulate this function for t going from -1 to 2 in increments of 0.1. Graph your results.
6. Let $x(t)$ be a pulse of unit intensity and unit duration,

$$\begin{aligned} x(t) &= 1, \quad \text{if} \quad 0 \leqslant t < 1, \\ &= 0, \quad \text{if} \quad t < 0 \text{ or } t \geqslant 1. \end{aligned}$$

Clearly $x(t) = \mu(t) - \mu(t-1)$. By linearity and invariance,

$$Gx(t) = G\mu(t) - G\mu(t-1). \tag{21}$$

Use this equation and the table of $G\mu(t)$ constructed in Exercise 5 to construct a table of $Gx(t)$ for t going from 0 to 2 in increments of 0.1. Graph your results. This graph will resemble Fig. 1 on p. 120 of Solomon's and Corbit's article.

Lecture 11. *Frequency Response and Fourier Transform*

We have seen that invariant linear systems map complex sinusoids into complex sinusoids of the same complex frequency. The only difference between input and output sinusoids is a multiplicative factor, $G(s)$, the transfer function of the system. Symbolically,

$$Ge^{st} = G(s)\, e^{st}. \tag{1}$$

This holds for any complex frequency, $s = \sigma + i\omega$, for which the integral

$$G(s) = \mathscr{L}\{g\}(s) = \int_{-\infty}^{\infty} e^{-st} g(t)\, dt \tag{2}$$

makes sense. If $\sigma = 0$, the integrand has absolute value

$$|e^{-i\omega t} g(t)| = |g(t)|,$$

and the integral is well defined whenever

$$\int_{-\infty}^{\infty} |g(t)|\, dt < \infty. \tag{3}$$

This condition is satisfied if, for example, G is nonanticipating and stable, and $G(s)$ is rational. We shall assume throughout this lecture that (3) is satisfied.

Taking $\sigma = 0$ in (1), we obtain

$$Ge^{i\omega t} = G(i\omega)\, e^{i\omega t}. \tag{4}$$

The nice thing about (4) is that it involves the "ordinary" frequency, ω, which is a bit more intuitive than the general complex frequency, s. The function $G(i\omega)$ of ω is termed the *frequency response function* of the system G. The full significance of this function becomes clearer if we write it in polar form

$$G(i\omega) = \rho(\omega)\, e^{i\varphi(\omega)}, \tag{5}$$

where

$$\rho(\omega) = |G(i\omega)| \tag{6}$$

and

$$\varphi(\omega) = \arg G(i\omega). \tag{7}$$

Plugging (5) into (4), we see that

$$Ge^{i\omega t} = \rho(\omega)\, e^{i(\omega t + \varphi(\omega))}.$$

The imaginary part of this equation is

$$G \sin \omega t = \rho(\omega) \sin(\omega t + \varphi(\omega)), \tag{8}$$

which brings us back into the realm of real sinusoidal inputs and outputs. Equation

(8) shows that the amplitude of an input sinusoid is altered by the factor $\rho(\omega)$, and the phase is altered by $\varphi(\omega)$. The functions $\rho(\omega)$ and $\varphi(\omega)$ are called the *gain* and *phase shift*, respectively. Graphs of $\varphi(\omega)$ and $\log \rho(\omega)$ against $\log \omega$ are called *Bode diagrams*. Use of $\log \omega$ presupposes $\omega > 0$. It is not difficult to show that $\rho(-\omega) = \rho(\omega)$ and $\varphi(-\omega) = -\varphi(\omega)$. Thus negative frequencies need not be considered.

An unknown transfer function, or, equivalently, its Bode diagram, can be determined experimentally. One simply compares input and output sinusoids for a number of different frequencies, ω, and interprets the results in the light of (8).

One small point: Eq. (7) only determines the phase shift up to an additive integer multiple of 2π, and, as far as (7) is concerned, this multiple can be different for different values of ω. In other words, if $n(\omega)$ is any integer-valued function of ω, then

$$\tilde{\varphi}(\omega) = \varphi(\omega) + 2\pi n(\omega) \tag{9}$$

is just as good as $\varphi(\omega)$ from the viewpoint of (7). Such nonsense is ruled out by requiring that the phase shift depend continuously on ω. If φ and $\tilde{\varphi}$ in (9) are continuous, then $n(\omega)$ must also be continuous. But the only continuous integer-valued functions are constants, $n(\omega) = n_0$ for some n_0. Then (9) reduces to

$$\tilde{\varphi}(\omega) = \varphi(\omega) + 2\pi n_0.$$

This is the only admissible transformation of the phase shift.

For $\omega = 0$, $e^{i\omega t} = 1$ is a constant (dc) input. Taking $s = 0$ in (2), we see that the corresponding gain and phase shift are

$$\rho(0) = \left| \int_{-\infty}^{\infty} g(t)\, dt \right| \tag{10}$$

and

$$\varphi(0) = \arg \int_{-\infty}^{\infty} g(t)\, dt,$$

or

$$\begin{aligned} \varphi(0) &= 0, \qquad &&\text{if } \int_{-\infty}^{\infty} g(t)\, dt > 0, \\ &= \pm\, \pi, \qquad &&\text{if } \int_{-\infty}^{\infty} g(t)\, dt < 0. \end{aligned} \tag{11}$$

Thus the integral of g can be inferred from the dc point (or dc limit) of the Bode diagram.

Examples

1. *n-stage RC filter.* For this system,

$$G(s) = (bs + 1)^{-n},$$

so

$$G(i\omega) = (1 + ib\omega)^{-n}, \tag{12}$$

and

$$\rho(\omega) = (1 + b^2\omega^2)^{-n/2}. \tag{13}$$

Clearly $\rho(\omega) \downarrow 0$ as $\omega \uparrow \infty$. This is a "low-pass filter." Note that

$$1 + b^2\omega^2 = b^2\omega^2(b^{-2}\omega^{-2} + 1),$$

so

$$\rho(\omega) = b^{-n}\omega^{-n}(b^{-2}\omega^{-2} + 1)^{-n/2}$$

and

$$\log \rho(\omega) = -n \log b - n \log \omega - (n/2) \log(b^{-2}\omega^{-2} + 1).$$

As $\omega \to \infty$, the final logarithm on the right approaches $\log 1 = 0$, hence

$$\log \rho(\omega) \to -n \log b - n \log \omega.$$

Thus the Bode diagram of ρ (see Fig. 28) approaches the straight line with slope $-n$ and y intercept $-n \log b$ for large frequencies.

To obtain the phase shift, we multiply numerator and denominator in (12) by the complex conjugate of the denominator. This yields

$$G(i\omega) = \frac{(1 - ib\omega)^n}{(1 + b^2\omega^2)^n}.$$

The denominator is real and positive, so it has no effect on the argument, $\varphi(\omega)$, of $G(i\omega)$. Thus

$$\varphi(\omega) = \arg(1 - ib\omega)^n$$

or

$$\varphi(\omega) = n \arg(1 - ib\omega). \tag{14}$$

As Fig. 29 shows, the argument of $1 - ib\omega$ goes from 0 to $-\pi/2$, hence $\varphi(\omega)$ goes from 0 to $-n\pi/2$, as ω goes from 0 to ∞. Since $\varphi(\omega) < 0$, the output may be

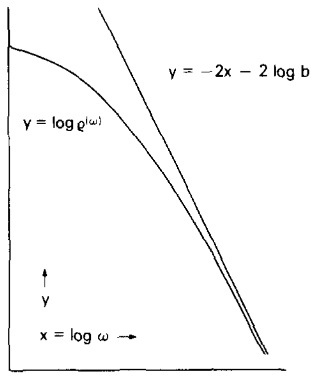

FIG. 28. Graph of $\log \rho(\omega)$ versus $\log \omega$ for a 2-stage *RC* filter.

regarded as lagging behind the input ($\omega t + \varphi(\omega)$ achieves the value 5, say, *later* than ωt). The lag increases with frequency, as can be seen from Fig. 29.

Combining our findings concerning $\rho(\omega)$ and $\varphi(\omega)$, we may plot the path of $G(i\omega)$ in the complex plane as ω goes from 0 to ∞. Such a plot is called a *Nyquist diagram.* Figure 30 is a Nyquist diagram for a three-stage *RC* filter. Arrows point in the direction of increasing ω.

2. *RCL circuit.* For this system,

$$G(s) = (as^2 + bs + 1)^{-1}.$$

We shall restrict our attention to the oscillatory case, in which

$$d = b^2 - 4a < 0.$$

Then

$$G(s) = [a(s - s_+)(s - s_-)]^{-1},$$

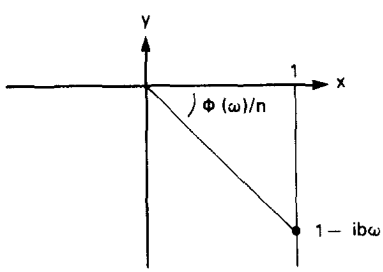

FIG. 29. Phase lag increases with frequency.

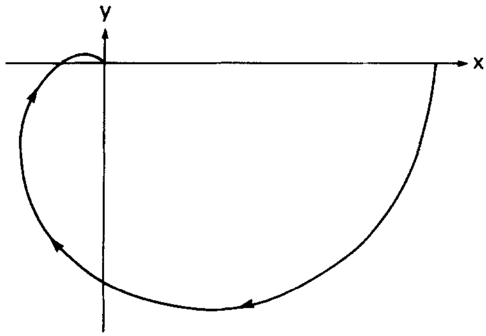

FIG. 30. Nyquist diagram for a 3-stage RC filter.

where

$$s_{\pm} = -\sigma_0 \pm i\omega_0,$$
$$\sigma_0 = b/(2a),$$

and

$$\omega_0 = \sqrt{|d|}/(2a).$$

It follows that

$$G(i\omega) = [a(i(\omega - \omega_0) + \sigma_0)(i(\omega + \omega_0) + \sigma_0)]^{-1},$$

hence

$$\rho(\omega) = [a^2(\sigma_0^2 + (\omega - \omega_0)^2)(\sigma_0^2 + (\omega + \omega_0)^2)]^{-1/2}.$$

The following theorem describes the main features of this function for $\omega \geqslant 0$.

THEOREM. *Suppose that $d < 0$. If $\omega_0 \leqslant \sigma_0$ then $\rho(\omega)$ is a decreasing function. If $\omega_0 > \sigma_0$, then $\rho(\omega)$ rises to a maximum at $\omega_1 = (\omega_0^2 - \sigma_0^2)^{1/2}$ and decreases thereafter. In either case, $\rho(\omega) \to 0$ as $\omega \to \infty$.*

To prove the theorem, one can take the natural logarithm of $\rho(\omega)$, differentiate, and consider whether the resulting expression is zero for any positive ω. I shall omit details.

The second case described in the theorem is the phenomenon of *resonance*, illustrated in Fig. 31. Clearly

$$\begin{aligned} \omega_1 &= (\omega_0^2 - \sigma_0^2)^{1/2} \\ &< (\omega_0^2)^{1/2} \\ &= \omega_0. \end{aligned}$$

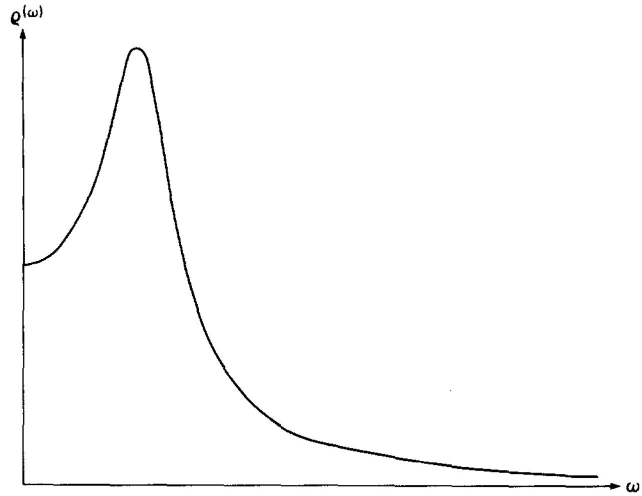

FIG. 31. Gain in an RCL circuit with $\omega_0 > \sigma_0$.

Thus the frequency, ω_1, of maximal responsiveness is less than the frequency, ω_0, of free oscillation after an impulse. (The impulse response, in our present notation, is

$$g(t) = (a\omega_0)^{-1} e^{-\sigma_0 t} \sin \omega_0 t$$

for $t \geqslant 0$.)

3. *Integrator.* The input and output of the integrator are supposed to be related by $y' = x$. Taking $x = e^{st}$ and $y = G(s) e^{st}$ in this equation yields $G(s) = s^{-1}$, as we noted in Lecture 7. This derivation applies to any nonzero s. For $s = i\omega$, we obtain

$$G(i\omega) = (i\omega)^{-1} = -i\omega^{-1},$$

hence

$$\rho(\omega) = \omega^{-1}$$

and

$$\varphi(\omega) = -\pi/2$$

for $\omega > 0$.

It is harder to interpret these functions than it was to derive them. The impulse response of the integrator is the unit step function, $g = \mu$. This function is unstable, condition (3) fails, and the integral in (2) is not well defined for $s = i\omega$ (since

$$\int_{-\infty}^{\infty} e^{-i\omega t} g(t)\, dt = \int_{0}^{\infty} e^{-i\omega t}\, dt$$
$$= \lim_{T\to\infty} \int_{0}^{T} e^{-i\omega t}\, dt$$
$$= \lim_{T\to\infty} (i\omega)^{-1} (1 - e^{-i\omega T}),$$

and the limit does not exist). By the same token, the familiar recipe

$$y(t) = \int_{-\infty}^{t} g(t-u)\, e^{su}\, du \tag{15}$$

can't be used to calculate the response of the integrator to the input $x(u) = e^{su}$ when $\mathrm{re}\, s = 0$. However, (2) and (15) are perfectly all right if $\mathrm{re}\, s > 0$. Thus, for the integrator, the transfer function rests on firmer theoretical ground than the frequency response function.

Flicker Detection

Suppose that a human subject is asked to judge whether or not a light is flickering. The intensity of the light is

$$L_0(1 + m \sin \omega t),$$

where L_0 is the average intensity and m controls the depth of modulation. For each L_0 and ω, the experimenter determines the smallest value of m at which flicker is visible.

Sperling (1964) (see also Chap. XIV of Cornsweet, 1970) describes a simple model for this situation. It assumes that the oscillatory component,

$$x(t) = L_0 m \sin \omega t,$$

of the light is processed by a linear system, G, that produces an output

$$y(t) = Gx(t)$$
$$= L_0 m G \sin \omega t$$
$$= L_0 m \rho(\omega) \sin(\omega t + \varphi(\omega)).$$

The subject perceives flicker if and only if the amplitude, $L_0 m \rho(\omega)$, of $y(t)$ equals or exceeds a threshold, which we may set equal to 1. The smallest value of m at which flicker is visible satisfies

$$L_0 m \rho(\omega) = \text{threshold} = 1.$$

(Let me digress for a moment to note that, when data from flicker–fusion experiments are considered within this theoretical framework, it is clear that $\rho(\omega)$ depends on L_0 as well as ω. The more general model of Sperling & Sondhi (1968) describes the effect of variation in L_0. We shall regard L_0 as fixed throughout the present discussion.)

Since L_0 is prescribed by the experimenter and m can be measured, it is possible to infer $\rho(\omega)$ from such an experiment. However, the experiment appears to yield no information about $\varphi(\omega)$. This raises the question: Is there any information in φ beyond what is contained in ρ? In other words, does ρ determine φ, and thus $G(i\omega)$ and g? The following simple example shows that the answer is "no."

Let G_+ and G_- be nonanticipating invariant linear systems with transfer functions

$$G_\pm(s) = \frac{s \pm 2}{(s+1)(s+3)}.$$

Since

$$|i\omega \pm 2| = (\omega^2 + 4)^{1/2},$$

both systems have the same gain,

$$\rho(\omega) = \frac{(\omega^2+4)^{1/2}}{(\omega^2+1)^{1/2}(\omega^2+9)^{1/2}}.$$

However, the impulse responses, $g_+(t)$ and $g_-(t)$, are quite different. As a consequence of Theorem 1 of Lecture 9,

$$g_+(t) = \tfrac{1}{2}e^{-t} + \tfrac{1}{2}e^{-3t}$$

and

$$g_-(t) = -\tfrac{3}{2}e^{-t} + \tfrac{5}{2}e^{-3t}$$

for $t \geqslant 0$. These functions are graphed in Fig. 32.

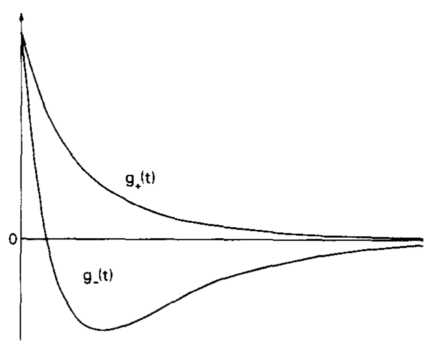

FIG. 32. Two impulse responses with the same gain.

Thus ρ doesn't determine φ, $G(i\omega)$, or g. Consequently, identification of an impulse response underlying visual detection requires consideration of inputs other than sinusoids. Roufs and Blommaert (in press) have recently derived an estimate of g from data on detectability of inputs consisting of two pulses of unequal magnitude.

Fourier Transform

Under rather general conditions, the entire Laplace transform theory can be specialized to the imaginary axis, $s = i\omega$. If

$$\int_{-\infty}^{\infty} |x(t)|\, dt < \infty, \tag{16}$$

then

$$X(i\omega) = \mathscr{L}\{x\}(i\omega) = \int_{-\infty}^{\infty} e^{-i\omega t} x(t)\, dt \tag{17}$$

makes sense. This quantity is called the *Fourier transform* of x. The meaning of the Fourier transform, like that of the more general Laplace transform, is contained in the inversion formula

$$x(t) = (2\pi)^{-1} \int_{-\infty}^{\infty} e^{i\omega t} X(i\omega)\, d\omega, \tag{18}$$

which is valid if $x'(u)$ exists and is continuous in some interval, $t - \varepsilon < u < t + \varepsilon$, around t. Thus $X(i\omega)$ describes "how much $e^{i\omega t}$ is present in $x(t)$." Taking $s = i\omega$ in (2), we see that *the frequency response function, $G(i\omega)$, is the Fourier transform of the impulse response.* Similarly,

$$Y(s) = G(s)\, X(s)$$

yields

$$Y(i\omega) = G(i\omega)\, X(i\omega), \tag{19}$$

which says that *the Fourier transform of the output is the Fourier transform of the input times the frequency response.*

These aspects of the "Fourier theory" are just special cases of the "Laplace theory." However, the following important theorem is special to Fourier analysis.

PARSEVAL'S THEOREM. *Suppose that $x(t)$ satisfies both* (16) *and*

$$\int_{-\infty}^{\infty} x(t)^2\, dt < \infty. \tag{20}$$

Then

$$\int_{-\infty}^{\infty} x(t)^2\, dt = (2\pi)^{-1} \int_{-\infty}^{\infty} |X(i\omega)|^2\, d\omega. \tag{21}$$

A proof of Parseval's Theorem is given on pages 149 and 150 of Schwarz and Friedland (1965).

The integral on the left in (21) and its square root are the most useful measures of the total magnitude of the function $x(t)$. In some physical applications, the integral is proportional to the energy associated with $x(t)$, and we may use the energy terminology metaphorically in the general theory. Equation (21) shows how the energy of $x(t)$ can be apportioned to different component frequencies. The energy corresponding to a small band of frequencies near ω is

$$(2\pi)^{-1}\, |X(i\omega)|^2\, d\omega.$$

Here $d\omega$ is the width of the. band expressed in radians, and $(2\pi)^{-1}\, d\omega$ is the bandwidth in cycles. Dividing out this factor, we see that $|X(i\omega)|^2$ is the energy per unit bandwith or *energy density*.

Taking absolute values in (19) and squaring, we find that

$$|Y(i\omega)|^2 = \rho(\omega)^2\, |X(i\omega)|^2. \tag{22}$$

The energy density of the output is the energy density of the input times the squared gain of the system.

APPENDIX. PROOF OF THE OSCILLATION THEOREM

This appendix gives a proof of Theorem 4 of Lecture 10.

According to Eq. (23) of Lecture 9,

$$g(t) = \mu(t) \sum_{j=1}^{\nu} P_j(t) \exp(r_j t). \tag{1}$$

All of the roots, r_j, of $Q(s)$ are assumed to be real. The polynomial $P_j(t)$ has degree *at most* $n_j - 1$, where $n_j \geqslant 1$ is the multiplicity of r_j. We may assume, without loss of generality, that $P(s)$ and $Q(s)$ have no common roots, in which case the degree of $P_j(t)$ is *exactly* $n_j - 1$. (To say that the degree of P_j is zero means that $P_j(t) = c$ for all t, for some *nonzero* constant c.) The sum of the multiplicities, n_j, is the degree, n, of $Q(s)$,

$$n = \sum_{j=1}^{\nu} n_j.$$

We wish to prove that $g(t)$ has fewer than n zeros for $t \geqslant 0$. Since every such zero is a zero of

$$h(t) = \sum_{j=1}^{\nu} P_j(t) \exp(r_j t), \tag{2}$$

it suffices to prove that $h(t)$ has fewer than n zeros for $-\infty < t < \infty$. At this point it is convenient to forget about impulse response functions, and to consider an arbitrary function, $h(t)$, defined by (2), where $P_j(t)$ is a polynomial. Let n_j be defined by

$$n_j = \text{degree } P_j + 1 \tag{3}$$

or, equivalently,

$$\text{degree } P_j = n_j - 1. \tag{4}$$

Finally, let

$$n_h = \sum_{j=1}^{\nu} n_j, \tag{5}$$

where the subscript "h" indicates that n_h depends on h. We wish to show that h has fewer than n_h zeros. Since n_h can be any positive integer, we must show that each of the following statements is correct:

S_n("statement n"): If $n_h = n$, then h has fewer than n zeros.

We must show that S_1 is true, S_2 is true, S_3 is true, etc.

We shall use the technique of mathematical induction. (We used this approach previously in the proof of Theorem 2 of Lecture 9.) For an inductive proof, it suffices to establish two things:

(A) S_1 is true.

(B) For any n, truth of S_n implies truth of S_{n+1}.

Proof of (A). This is easy. If $n_h = 1$, it follows from (5) and $n_j \geqslant 1$ that $\nu = 1$ and $n_1 = 1$. Thus, by (2),

$$h(t) = P_1(t) \exp(r_1 t),$$

and, by (4), degree $P_1 = 1 - 1 = 0$. Thus $P_1(t) = c \neq 0$ for all t, and $h(t)$ has *no* zeros, as was to be shown.

Proof of (B). Suppose that S_n is true for some positive integer, n, and let $n_h = n + 1$. We shall show that $h(t)$ has at most n zeros. *We may, and shall, assume that* $r_\nu = 0$. For $h(t)$ has the same zeros as

$$\begin{aligned} h^*(t) &= h(t) \exp(-r_\nu t) \\ &= \sum_{j=1}^{\nu} P_j(t) \exp((r_j - r_\nu)t), \end{aligned}$$

and the νth exponential constant for this function is $r_\nu - r_\nu = 0$.

Differentiating (2), we obtain

$$h'(t) = \sum_{j=1}^{\nu} Q_j(t) \exp(r_j t), \tag{6}$$

where $Q_j(t)$ is the polynomial

$$Q_j(t) = P_j'(t) + r_j P_j(t).$$

The r_j's are all different, and $r_\nu = 0$, so $r_j \neq 0$ for $j < \nu$. It follows easily from this that

$$\text{degree } Q_j = \text{degree } P_j, \qquad j < \nu. \tag{7}$$

Also

$$Q_\nu(t) = P_\nu'(t).$$

Thus, if degree $P_\nu \geqslant 1$, then

$$\text{degree } Q_\nu = \text{degree } P_\nu - 1. \tag{8}$$

Comparing (2) and (6), and using (3), (5), (7), and (8), we conclude that

$$\begin{aligned} n_{h'} &= n_h - 1 \\ &= (n+1) - 1, \end{aligned}$$

or

$$n_{h'} = n. \tag{9}$$

This is also valid if degree $P_\nu = 0$, for, in that case, $Q_\nu(t) = 0$ for all t and

$$h'(t) = \sum_{j=1}^{\nu-1} Q_j(t) \exp(r_j t).$$

Thus (9) is valid unconditionally. *It then follows from S_n, which we are assuming, that h' has fewer than n zeros.*

Our objective is to show that h has fewer than $n+1$ zeros. We shall do this by assuming the contrary and obtaining a contradiction. Between any two zeros of h there is a zero of h'. (This follows from the "mean value theorem of differential calculus.") Hence, if h has $n+1$ or more zeros, then h' has n or more zeros, contrary to our conclusion in the previous paragraph. This contradiction establishes that h has fewer than $n+1$ zeros.

This completes the proof of (B), which was all that remained of the proof that $h(t)$ has fewer than n_h zeros for $-\infty < t < \infty$. As we have noted, this implies that $g(t)$ has fewer than n zeros for $t \geqslant 0$, where $n = \text{degree } Q$.

The comparable assertions concerning $g'(t)$, $g''(t)$, etc., follow from what has already been done, since $h'(t)$, $h''(t)$, etc., are "just like $h(t)$" (compare, for example, (6) and (2)).

This completes the proof of Theorem 4 of Lecture 10.

References

BRACEWELL, R. N. *The Fourier transform and its applications* (2nd ed.). New York: McGraw–Hill, 1978.

CORNSWEET, T. N. *Visual perception*. New York: Academic Press, 1970.

FEYNMAN, R. P., LEIGHTON, R. B., & SANDS, M. *The Feynman lectures on physics* (Vol. I). Reading, Mass.: Addison–Wesley, 1963.

HARRIS, C. S. (Ed.) *Visual coding and adaptability*. Hillsdale, N. J.: Lawrence Erlbaum, 1980.

HILL, A. V. Excitation and accommodation in nerve. *Proceedings of the Royal Society, Series B*, 1936, **119**, 305–355.

KNOPP, K. *Elements of the theory of functions*. New York: Dover, 1952.

LICKLIDER, J. C. R. Quasi-linear operator models in the study of manual tracking. In R. D. Luce (Ed.), *Developments in mathematical psychology*. Glencoe, Ill.: Free Press, 1960. Pp. 167–279.

MCFARLAND, D. J. *Feedback mechanisms in animal behaviour*. London: Academic Press, 1971.

MILSUM, J. H. *Biological control system analysis*. New York: McGraw–Hill, 1966.

ROUFS, J. A. J., & BLOMMAERT, F. J. J. Temporal impulse and step responses of the human eye obtained psychophysically by means of a drift-correcting perturbation technique. *Vision Research*, in press.

SCHWARZ, R. J., & FRIEDLAND, B. *Linear systems*. New York: McGraw–Hill, 1965.

SOLOMON, R. L., & CORBIT, J. D. An opponent-process theory of motivation: I. Temporal dynamics of affect. *Psychological Review*, 1974, **81**, 119–145.

SPERLING, G. Linear theory and the psychophysics of flicker. *Documenta Ophthalmologica*, 1964, **18**, 3–15.

SPERLING, G., & SONDHI, M. M. Model for visual luminance discrimination and flicker detection. *Journal of the Optical Society of America*, 1968, **58**, 1133–1145.

STARK, L. *Neurological control systems*. New York: Plenum, 1968.

RECEIVED: August 30, 1980

Index

Standard notations are given in brackets.