# Signal Distortion in Nonlinear Feedback Systems

By I. W. SANDBERG

(Manuscript received May 6, 1963)

*This paper reports on some properties of the solutions to the functional equation* $s_2(t) = \varphi[\mathbf{C}s_2(t) + s_1(t)]$, *where* $\varphi$ *is a nonlinear function, the operator* $\mathbf{C}$ *is a convolution, and* $s_1$ *is a known function belonging to a prescribed Banach space.*

*The equation plays a central role in the theory of signal transmission through a general physical system containing linear time-invariant elements and a single time-variable nonlinear element. In particular we establish conditions under which* $s_2(t)$ *is the fixed point of a contraction mapping of the Banach space into itself and we discuss some consequences of this result.*

*As a direct application, we consider the range of validity of two simple cascade flow graphs (i.e., flow graphs without feedback loops) for approximately determining the signal distortion in nonlinear feedback systems when the distortion is small. Our discussion is not restricted to specific types of nonlinear characteristics.*

## I. INTRODUCTION

This paper reports on some properties of the solutions to the functional equation $s_2(t) = \varphi[\mathbf{C}s_2(t) + s_1(t)]$, where $\varphi$ is a nonlinear function, the operator $\mathbf{C}$ is a convolution, and $s_1$ is a known function belonging to a prescribed Banach space.

The equation plays a central role in the theory of signal transmission through a general physical system containing linear time-invariant elements and a single time-variable nonlinear element. In particular we establish conditions under which $s_2(t)$ is the fixed point of a contraction mapping of the Banach space into itself and we discuss some consequences of this result.

As a direct application, we consider the range of validity of a simple cascade flow graph (i.e., a flow graph without feedback loops) for approximately determining the signal distortion in nonlinear feedback systems when the distortion is small. Our discussion is not restricted to specific types of nonlinear characteristics.

Except in cases in which the nonlinearity is very small, our results establish the utility of the graph only when (in a certain precise sense) the feedback around the nonlinear element is small. However, our results show that the range of validity of this flow graph is very much greater than that indicated by an earlier writer[1] who has considered this question for the case in which the nonlinear characteristic is of the form $x + \epsilon x^m$, where $\epsilon$ and $m$ are real constants with $m$ an odd positive integer.

As is well known, large amounts of feedback are often present in physical systems. In fact, large amounts of feedback are often used to reduce nonlinear distortion. Consequently the established range of validity of the flow graph mentioned above does not include the most important cases of interest. To deal with such situations we propose an alternative, but very closely related, flow graph for approximately determining the signal distortion when the distortion is small. It appears that the range of validity of this graph includes the vast majority of cases of engineering interest.

Section II considers some mathematical preliminaries. In Section III we describe a model of the physical system under consideration and show the relevance of the functional equation mentioned above. Section IV presents some preliminary results which are concerned with the properties of certain linear operators. In the remaining sections we consider both some properties of the solutions to the functional equation and engineering implications of the results.

## II. MATHEMATICAL PRELIMINARIES

Let $\mathcal{R} = [\Theta,\rho]$ be an arbitrary metric space.[2] A mapping $\mathbf{A}$ of the space $\mathcal{R}$ into itself is said to be a contraction if there exists a number $k < 1$ such that

$$\rho(\mathbf{A}x,\mathbf{A}y) \leqq k\rho(x,y)$$

for any two elements $x,y \in \Theta$. The contraction-mapping fixed-point theorem[2] is basic to much of the subsequent discussion. It states that every contraction mapping defined in a complete metric space $\mathcal{R}$ has one and only one fixed point (i.e., there exists a unique element $z \in \Theta$ such that $\mathbf{A}z = z$). Furthermore $z = \lim_{n\to\infty} \mathbf{A}^n x_0$, where $x_0$ is an arbitrary element of $\Theta$.

Throughout the discussion $\mathcal{L}_2$ denotes the space of complex-valued square-integrable functions defined on the real interval $(-\infty, \infty)$. The norm of $f(t) \in \mathfrak{L}_2$ is denoted by $\| f \|_2$ and is defined by

$$\| f \|_2^2 = \int_{-\infty}^{\infty} | f(t) |^2 \, dt.$$

The symbol $\mathfrak{L}_1$ denotes the space of absolutely integrable functions defined on $(-\infty, \infty)$. We shall use the symbols $\mathfrak{L}_{2R}$ and $\mathfrak{L}_{1R}$, respectively, to denote the intersections of the spaces $\mathfrak{L}_2$ and $\mathfrak{L}_1$ with the set of real-valued functions. It is well known that $\mathfrak{L}_2$ and $\mathfrak{L}_{2R}$ are Banach spaces.

We take as the definition of the Fourier transform of $f(t) \in \mathfrak{L}_2 \cup \mathfrak{L}_1$

$$F(\omega) = \int_{-\infty}^{\infty} f(t) \, e^{-i\omega t} \, dt, \qquad f \in \mathfrak{L}_1$$

$$F(\omega) = \text{l.i.m.} \int_{-\infty}^{\infty} f(t) \, e^{-i\omega t} \, dt, \qquad f \in (\mathfrak{L}_2 - \mathfrak{L}_1)$$

and consequently when $f(t) \in \mathfrak{L}_2$

$$f(t) = \text{l.i.m.} \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) \, e^{i\omega t} \, d\omega.$$

With this definition, the Plancherel identity reads

$$2\pi \int_{-\infty}^{\infty} f(t)\underset{\sim}{g}(t) \, dt = \int_{-\infty}^{\infty} F(\omega)\underset{\sim}{G}(\omega) \, d\omega, \qquad f, g \in \mathfrak{L}_2 .$$

Throughout the discussion $\mathcal{K}(\Sigma)$ denotes the space of bounded real-valued functions that $(i)$ are defined on the real interval $(-\infty, \infty)$ and $(ii)$ are continuous on $(-\infty, \infty) - \Sigma$, where $\Sigma$ is an arbitrary fixed finite or infinite set of isolated points.* The norm of $f \in \mathcal{K}(\Sigma)$ is denoted by $\| f \|_\infty$ and is defined by

$$\| f \|_\infty = \sup_t | f(t) | .$$

With this norm $\mathcal{K}(\Sigma)$ is a Banach space.† The norm of a linear operator $\mathbf{Q}$ defined on $\mathcal{K}(\Sigma)$ is denoted by $\| \mathbf{Q} \|_\infty$ and similarly for the norm of a linear operator defined on $\mathfrak{L}_2$.

We shall say that a real-valued function $f(t)$ belongs to $\mathfrak{D}$ if and only if there exists a function $\hat{f}(t)$ that agrees with $f(t)$ almost everywhere

* Various signals of interest in communication systems such as pulses are not contained in $\mathcal{K}(\Sigma)$ if $\Sigma$ is the null set.

† Any Cauchy sequence of functions belonging to $\mathcal{K}(\Sigma)$ converges to a bounded continuous function on $(-\infty, \infty) - \Sigma$ (since the sequence converges uniformly on $(-\infty, \infty) - \Sigma$). Since, in addition, the sequence converges at each point of discontinuity, $\mathcal{K}(\Sigma)$ is complete.

and is such that the set of points at which* sign $[\hat{f}(t)]$ is discontinuous is a set of isolated points.

The symbols **I** and **O** are used throughout to denote, respectively, the identity operator and the null operator (i.e., for all $f$, $\mathbf{O}f = 0$).

## III. MATHEMATICAL MODEL OF THE PHYSICAL SYSTEM

Consider a physical system containing linear time-invariant elements and a single time-variable nonlinear element. Let $s_1$ and $s_2$, respectively, denote the system's input and output signals and let $v$ and $w$, respectively, denote the input and output signals associated with the nonlinear device, which is assumed to be characterized by the equation

$$w = \varphi(v,t) = \varphi[v] \tag{1}$$

in which $\varphi(v,t)$ is a real-valued function of the real variables $v$ and $t$.

We shall consider separately two cases:

($i$) $s_1, s_2, v, w \in \mathcal{K}(\Sigma)$ for some $\Sigma$

($ii$) $s_1, s_2, v, w \in \mathcal{L}_{2R}$.

It is assumed in each case that there exist well-defined linear operators $\mathbf{\Gamma}$ and $\mathbf{\Lambda}$ such that† $v = \mathbf{\Gamma}[s_1, w]$ and $s_2 = \mathbf{\Lambda}[s_1, w]$. It is convenient to define four linear operators **A**, **B**, **C**, and **D** in the following manner

$$\begin{aligned} v &= \mathbf{\Gamma}[s_1, w] = \mathbf{\Gamma}[s_1, 0] + \mathbf{\Gamma}[0, w] \\ &= \mathbf{A}s_1 + \mathbf{C}w \end{aligned} \tag{2}$$

$$\begin{aligned} s_2 &= \mathbf{\Lambda}[s_1, w] = \mathbf{\Lambda}[s_1, 0] + \mathbf{\Lambda}[0, w] \\ &= \mathbf{D}s_1 + \mathbf{B}w. \end{aligned} \tag{3}$$

The relation between $s_1$, $s_2$, $v$, and $w$ is summarized by the flow graph in Fig. 1.

As a very simple illustration of the generality of the graph in Fig. 1, observe that the flow graph of the classical single loop feedback system in Fig. 2(a) in which **E** and **F** are linear operators can readily be reduced to the form shown in Fig. 1. The reduced graph is given in Fig. 2(b).

Our concern is with the influence of the nonlinear element represented by $\varphi$. Hence it is sufficient to consider the situation in Fig. 1 in which $\mathbf{D} = \mathbf{O}$ and $\mathbf{A} = \mathbf{B} = \mathbf{I}$. The corresponding graph is shown in Fig. 3. For this graph, using (1), (2), and (3)

$$s_2 = \varphi[\mathbf{C}s_2 + s_1]. \tag{4}$$

* Let sign $[\hat{f}(t)] = 1$ when $\hat{f}(t) = 0$.

† This is essentially the same model used by the writer in another study.[3]

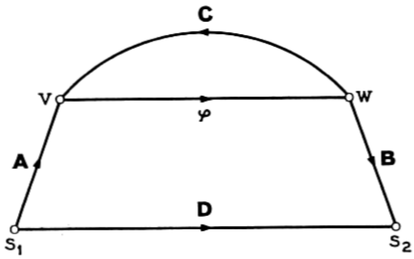


Fig. 1 — Flow-graph representation of a general transmission system containing linear elements and a single time-variable nonlinear element $\varphi$.

### 3.1 *The Time-Variable Nonlinear Element and Definition of Signal Distortion*

It is assumed throughout that $\varphi[f(t)]$ is measurable whenever $f$ is measurable, that $\varphi(0,t) = 0$ for all $t$, and that for all $t$ and all $v_1 \geqq v_2$

$$\alpha(v_1 - v_2) \leqq \varphi(v_1, t) - \varphi(v_2, t) \leqq \beta(v_1 - v_2)$$

where $\alpha$ and $\beta$ are real constants.

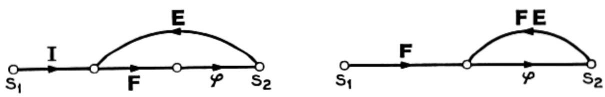


Fig. 2 — Two flow graphs with identical transmission from $s_1$ to $s_2$.

In our application of the theorems in Sections V and VII, we shall suppose that $\varphi(v,t) = v + \tilde{\varphi}(v,t)$ where, for all $t$, $\tilde{\varphi}(v,t)$ is of order less than $v$ as $v \to 0$. That is,* we shall suppose that for sufficiently small

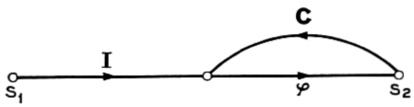


Fig. 3 — Basic flow graph for studying the influence of $\varphi$.

input signals the element represented by $\varphi$ behaves essentially as a unit gain amplifier and hence that, for such signals, the system in Fig. 1 behaves essentially as a linear time-invariant system. Let $s_{02}$ denote the

* Note that in the frequently encountered case in which $\varphi$ is independent of $t$ and $\varphi'$ is a monotone decreasing function of $v$, $\beta = 1$.

output signal $s_2$ in Fig. 1 when $\varphi(v,t) = v$. We shall say that $(s_2 - s_{02})$ is the signal distortion introduced by the departure of $\varphi(v,t)$ from $v$.

### 3.2 *The Operator* **C**

Unless stated otherwise, it is assumed that

$$\mathbf{C}f = \int_{-\infty}^{\infty} c(t - \tau) f(\tau)\, d\tau \tag{5}$$

where $c(t)$ is a real-valued function of $t$. In cases of engineering interest **C** is a causal (i.e., $c(t) = 0$, $t < 0$). However our mathematical results are not restricted to cases in which **C** is causal.

## IV. PRELIMINARY RESULTS

This section is concerned with a proof of

*Theorem I*: (*a*) *Let* $c(t) \in \mathcal{L}_{1R}$ *and* $C(\omega) \neq 1$. *Then* $(\mathbf{I} - \mathbf{C})$ *is a bounded mapping of* $\mathcal{K}(\Sigma)$ *into itself that possesses a bounded inverse. In fact, there exists a function* $h(t) \in \mathcal{L}_{1R}$, *with Fourier transform* $C(\omega)[1 - C(\omega)]^{-1}$, *such that*

$$(\mathbf{I} - \mathbf{C})^{-1} g = g + \int_{-\infty}^{\infty} h(t - \tau) g(\tau)\, d\tau$$

*for any* $g \in \mathcal{K}(\Sigma)$. *If* $h(t) \in \mathfrak{D}$,

$$\|(\mathbf{I} - \mathbf{C})^{-1}\|_\infty = 1 + \int_{-\infty}^{\infty} |\, h(t)\,|\, dt.$$

(*b*) *Suppose alternatively* that* $c(t) \in \mathcal{L}_{2R}$, *ess* $\sup_\omega |\, C(\omega)\,| < \infty$, *and that* $\inf_\omega |\, 1 - C(\omega)\,| > 0$. *Then* $(\mathbf{I} - \mathbf{C})$ *is a bounded mapping of* $\mathcal{L}_{2R}$ *into itself that possesses a bounded inverse. Moreover*

$$\|(\mathbf{I} - \mathbf{C})^{-1}\|_2 = \text{ess}\sup_\omega |\, 1 - C(\omega)\,|^{-1}$$

$$\|\mathbf{C}\|_2 = \text{ess}\sup_\omega |\, C(\omega)\,|.$$

### 4.1 *Proof of Part* (*a*)

Since $c(t) \in \mathcal{L}_{1R}$, the validity of the assertion that $(\mathbf{I} - \mathbf{C})$ is a bounded

* The notation ess $\sup_\omega Q(\omega)$ denotes $\inf_{\mathfrak{N}} \sup_{\omega \notin \mathfrak{N}} Q(\omega)$ where $\mathfrak{N}$ is an arbitrary zero-measure subset of the real line. In at least almost all cases of engineering interest, the "essential supremum" of the modulus of a Fourier transform is equal to its supremum.

mapping of $\mathcal{K}(\Sigma)$ into itself is obvious. For the remainder of part (a) we need

*Lemma I: Let* $c(t) \varepsilon \mathcal{L}_{1R}$ *and* $C(\omega) \neq 1$. *Then there exists a function* $h(t) \varepsilon \mathcal{L}_{1R}$, *with Fourier transform* $C(\omega)\,[1 - C(\omega)]^{-1}$, *such that*

$$h(t) - c(t) = \int_{-\infty}^{\infty} h(t - \tau)c(\tau)\,d\tau$$

*almost everywhere.*

4.1.1 *Proof of Lemma I:*

It is known[4]* that if $c(t) \varepsilon \mathcal{L}_{1R}$ and $\inf_{\omega} |\, 1 - C(\omega) \,| > 0$, there exists an $f(t)$ of bounded total variation on $(-\infty, \infty)$ such that

$$[1 - C(\omega)]^{-1} = \int_{-\infty}^{\infty} e^{-i\omega t}\,df(t).$$

Under these conditions, it follows that

$$h(t) = \int_{-\infty}^{\infty} c(t - \tau)\,df(\tau)$$

is an element of $\mathcal{L}_{1R}$ which possesses the required Fourier transform. However, since $C(\omega)$ is uniformly continuous and $C(\omega) \to 0$ as $\omega \to \infty$, the inequality $\inf_{\omega} |\, 1 - C(\omega) \,| > 0$ is satisfied if and only if $C(\omega) \neq 1$. Thus the assumptions in Lemma I imply the existence of a function $h(t) \varepsilon \mathcal{L}_{1R}$ with the stated transform.† Since

$$\frac{C(\omega)}{1 - C(\omega)} - C(\omega) = \frac{C(\omega)}{1 - C(\omega)}\,C(\omega)$$

the Fourier transforms of $[h(t) - c(t)]$ and

$$\int_{-\infty}^{\infty} h(t - \tau)c(\tau)\,d\tau$$

are equal. This establishes the equation stated in the lemma.

Let $g(t)$ denote any element of $\mathcal{K}(\Sigma)$ and assume that there exists an $f(t) \varepsilon \mathcal{K}(\Sigma)$ such that

---

* The writer is indebted to V. E. Beneš for directing attention to the result in Ref. 4.

† A moment's reflection will show that when $C(\omega)$ is rational in $\omega$, a proof of this result follows directly from the identification of the terms in its partial-fraction expansion.

$$g(t) = (\mathbf{I} - \mathbf{C})f = f(t) - \int_{-\infty}^{\infty} c(t-\tau)f(\tau)\,d\tau. \tag{6}$$

It is certainly true that (6) implies

$$\begin{aligned}&\int_{-\infty}^{\infty} h(t-\tau)g(\tau)\,d\tau \\ &= \int_{-\infty}^{\infty} h(t-\tau)f(\tau)\,d\tau - \int_{-\infty}^{\infty} h(t-\tau)\left[\int_{-\infty}^{\infty} c(\tau-u)f(u)\,du\right]d\tau.\end{aligned} \tag{7}$$

Since $f$ is bounded and $h,c \in \mathfrak{L}_{1R}$, Fubini's theorem implies that the last integral can be written as

$$\int_{-\infty}^{\infty} f(u)\left[\int_{-\infty}^{\infty} h(t-\tau)c(\tau-u)\,d\tau\right]du$$

and hence, in accordance with the lemma, as

$$\int_{-\infty}^{\infty} h(t-\tau)f(\tau)\,d\tau - \int_{-\infty}^{\infty} c(t-\tau)f(\tau)\,d\tau. \tag{8}$$

Therefore,

$$\int_{-\infty}^{\infty} h(t-\tau)g(\tau)\,d\tau = \int_{-\infty}^{\infty} c(t-\tau)f(\tau)\,d\tau = f(t) - g(t).$$

Thus, if there exists an $f \in \mathcal{K}(\Sigma)$ such that $(\mathbf{I} - \mathbf{C})f = g$,

$$f(t) = g(t) + \int_{-\infty}^{\infty} h(t-\tau)g(\tau)\,d\tau. \tag{9}$$

However, direct substitution and an application of Fubini's theorem show that the right-hand side of (9) is a solution of (6). Hence it is *the* solution.

Since $\| (\mathbf{I} - \mathbf{C})^{-1} \|_\infty = \sup\{ \| f \|_\infty : (\mathbf{I} - \mathbf{C})f = g;\ f,g \in \mathcal{K}(\Sigma);\ \| g \|_\infty = 1\}$, it is evident from (9) that

$$\| (\mathbf{I} - \mathbf{C})^{-1} \|_\infty \leqq 1 + \int_{-\infty}^{\infty} | h(t) |\,dt.$$

We shall next show that if $h(t) \in \mathfrak{D}$

$$\| (\mathbf{I} - \mathbf{C})^{-1} \|_\infty \geqq 1 + \int_{-\infty}^{\infty} | h(t) |\,dt - \delta,$$

where $\delta$ is an arbitrary positive number and hence that

$$\| (\mathbf{I} - \mathbf{C})^{-1} \|_\infty = 1 + \int_{-\infty}^{\infty} | h(t) |\,dt.$$

Choose a real number $t_0$ and consider

$$g(t_0) + \int_{-\infty}^{\infty} h(t_0 - \tau)g(\tau)\, d\tau. \tag{10}$$

We may assume that the set of points at which sign $[h(t)]$ is discontinuous is a set of isolated points. Let $\Xi$ denote the union of the closed intervals of length $\delta_1$ centered at $t_1$ and at each of the discontinuities of sign $[h(t_0 - \tau)]$ regarded as a function of $\tau$. Let $\Xi_*$ denote the complement of $\Xi$ with respect to the real line. For any $\delta > 0$, choose $\delta_1$ such that*

$$\int_{\Xi} | h(t_0 - \tau) |\, d\tau \leqq \tfrac{1}{2}\delta.$$

Choose $g(t) \in \mathcal{K}(\Sigma)$ such that $g(t_0) = 1$, $\| g \|_\infty = 1$, and

$$g(\tau) = \text{sign}\,[h(t_0 - \tau)], \qquad \tau \in \Xi_*$$

Then

$$g(t_0) + \int_{-\infty}^{\infty} h(t_0 - \tau)g(\tau)\, d\tau = 1 + \int_{-\infty}^{\infty} | h(t) |\, dt + \int_{\Xi} h(t_0 - \tau)\{g(\tau) - \text{sign}\,[h(t_0 - \tau)]\}\, d\tau,$$

and

$$\left| g(t_0) + \int_{-\infty}^{\infty} h(t_0 - \tau)g(\tau)\, d\tau \right| \geqq 1 + \int_{-\infty}^{\infty} | h(t) |\, dt - \delta.$$

This completes the proof of the first part of Theorem I.

Of course similar arguments show that if $c(t) \in \mathfrak{D}$,

$$\| \mathbf{C} \|_\infty = \int_{-\infty}^{\infty} | c(t) |\, dt.$$

4.2 *Proof of Part* $(b)$:

The proof of this part involves essentially the same arguments presented elsewhere.[3]

Let $f \in \mathfrak{L}_{2R}$. Then, using Plancherel's identity,

* The integral of $| h(t_0 - \tau) |$ over $\Xi$ does not exceed the sum of the integrals over $| \tau | \geqq T$ and $\Xi - [-T,T]_*$ where $[-T,T]_*$ denotes the complement of $[-T,T]$ with respect to the real line. The first integral can be made arbitrarily small by choosing $T$ sufficiently large and with fixed $T$ the second integral can be made arbitrarily small by choosing $\delta_1$ sufficiently small.

$$\begin{aligned}\| \mathbf{C}f \|_2^2 &= \frac{1}{2\pi}\int_{-\infty}^{\infty} | C(\omega)F(\omega) |^2 \, d\omega \\ &\leqq \operatorname*{ess\,sup}_{\omega} | C(\omega) |^2 \, \| f \|_2^2 .\end{aligned}$$

Thus $\mathbf{C}$, and hence $(\mathbf{I} - \mathbf{C})$, are bounded.

Now consider the equation

$$(\mathbf{I} - \mathbf{C})f = g; \qquad g \in \mathfrak{L}_{2R}, \qquad \| g \|_2 = 1.$$

Since $\inf_{\omega} | 1 - C(\omega) | > 0$, there exists a unique solution $f \in \mathfrak{L}_{2R}$ and $F(\omega) = G(\omega)[1 - C(\omega)]^{-1}$. Again using Plancherel's identity

$$\begin{aligned}\| f \|_2^2 = \| (\mathbf{I} - \mathbf{C})^{-1} g \|_2^2 &= \frac{1}{2\pi}\int_{-\infty}^{\infty} | 1 - C(\omega) |^{-2} \, | G(\omega) |^2 \, d\omega \\ &\leqq \operatorname*{ess\,sup}_{\omega} | 1 - C(\omega) |^{-2} \, \| g \|_2^2 .\end{aligned} \tag{11}$$

Clearly, $\| (\mathbf{I} - \mathbf{C})^{-1} \|_2 \leqq \operatorname*{ess\,sup}_{\omega} | 1 - C(\omega) |^{-1}$.

According to the definition of the essential supremum of a function, for any $\delta > 0$ there exists a set of values of $\omega$ of nonzero measure such that

$$| 1 - C(\omega) |^{-1} > \operatorname*{ess\,sup}_{\omega} | 1 - C(\omega) |^{-1} - \delta .$$

Since $| G(\omega) |$ is permitted to vanish only on the complement of such a set, it is evident from (11) that

$$\| (\mathbf{I} - \mathbf{C})^{-1} \|_2 \geqq \operatorname*{ess\,sup}_{\omega} | 1 - C(\omega) |^{-1} - \delta$$

for any $\delta > 0$. Thus, in view of the upper bound on $\| (\mathbf{I} - \mathbf{C})^{-1} \|_2$,

$$\| (\mathbf{I} - \mathbf{C})^{-1} \|_2 = \operatorname*{ess\,sup}_{\omega} | 1 - C(\omega) |^{-1} .$$

A similar argument shows that $\| \mathbf{C} \|_2 = \operatorname*{ess\,sup}_{\omega} | C(\omega) |$. This completes the proof of Theorem I.

## V. PROPERTIES OF SOLUTIONS TO $s_2 = \varphi[\mathbf{C}s_2 + s_1]$

*Theorem II: Let* $c(t) \in \mathfrak{L}_{1R}$, $C(\omega) \neq 1$, *and let* $\varphi[v] = v + \tilde{\varphi}[v]$ *be as defined in Section* 3.1. *Let* $\varphi[f]$ *be continuous with respect to* $t$ *on the complement of* $\Sigma$ *whenever* $f \in \mathcal{K}(\Sigma)$. *Suppose that*

$$\| (\mathbf{I} - \mathbf{C})^{-1} \|_\infty \cdot \| \mathbf{C} \|_\infty \, max \, (| 1 - \alpha |, | \beta - 1 |) = r < 1.$$

*Then for any* $s_1 \in \mathcal{K}(\Sigma)$, *there exists a unique* $s_2 \in \mathcal{K}(\Sigma)$ *such that* $s_2 =$

$\varphi[\mathbf{C}s_2 + s_1]$. *Furthermore* $s_2 = \lim_{n \to \infty} s_{2n}$ *where*

$$s_{2n} = (\mathbf{I} - \mathbf{C})^{-1}\tilde{\varphi}[\mathbf{C}s_{2(n-1)} + s_1] + (\mathbf{I} - \mathbf{C})^{-1}s_1$$

*and* $s_{20}$ *is an arbitrary element of* $\mathcal{K}(\Sigma)$. *The nth approximation* $s_{2n}$ *satisfies*

$$\| s_{2n} - s_2 \|_\infty \leqq \frac{r^n}{1 - r} \| s_{21} - s_{20} \|_\infty .$$

Before proceeding to the proof of Theorem II we state the analogous result for the space $\mathcal{L}_{2R}$.

*Theorem III*: *Let* $c(t) \in \mathcal{L}_{2R}$, $\sup_\omega | C(\omega) | < \infty$, $\inf_\omega | 1 - C(\omega) | > 0$, *and let* $\varphi[v] = v + \tilde{\varphi}[v]$ *be as defined in Section 3.1 Suppose that*

$$\| (\mathbf{I} - \mathbf{C})^{-1} \|_2 \cdot \| \mathbf{C} \|_2 \, max \, (|1 - \alpha |, | \beta - 1 |) = r < 1.$$

*Then the conclusion of Theorem II follows with* $\mathcal{K}(\Sigma)$ *replaced with* $\mathcal{L}_{2R}$ *and with the* $\mathcal{K}(\Sigma)$ *norm replaced with the* $\mathcal{L}_{2R}$ *norm.*

### 5.1 *Proof of Theorem II*

We have

$$s_2 = \mathbf{C}s_2 + s_1 + \tilde{\varphi}[\mathbf{C}s_2 + s_1]$$

and hence, in view of the first part of Theorem I and the assumption that $C(\omega) \neq 1$, $s_2 = \mathbf{L}s_2$ where

$$\mathbf{L}s_2 = (\mathbf{I} - \mathbf{C})^{-1}\tilde{\varphi}[\mathbf{C}s_2 + s_1] + (\mathbf{I} - \mathbf{C})^{-1}s_1 .$$

It is evident that $\mathbf{L}$ is a mapping of $\mathcal{K}(\Sigma)$ into itself. We shall show that under the conditions stated in the theorem $\mathbf{L}$ is in fact a contraction mapping of $\mathcal{K}(\Sigma)$ into itself. Let $f, g \in \mathcal{K}(\Sigma)$ and observe that

$$\begin{aligned} \| \mathbf{L}f - \mathbf{L}g \|_\infty &= \| (\mathbf{I} - \mathbf{C})^{-1}\{\tilde{\varphi}[\mathbf{C}f + s_1] - \tilde{\varphi}[\mathbf{C}g + s_1]\} \|_\infty \\ &\leqq \| (\mathbf{I} - \mathbf{C})^{-1} \|_\infty \cdot \| \tilde{\varphi}[\mathbf{C}f + s_1] - \tilde{\varphi}[\mathbf{C}g + s_1] \|_\infty \end{aligned}$$

and that

$$\begin{aligned} &\| \tilde{\varphi}[\mathbf{C}f + s_1] - \tilde{\varphi}[\mathbf{C}g + s_1] \|_\infty \\ &\qquad = \left\| \left( \frac{\varphi[\mathbf{C}f + s_1] - \varphi[\mathbf{C}g + s_1]}{\mathbf{C}f - \mathbf{C}g} - 1 \right) \mathbf{C}(f - g) \right\|_\infty \\ &\qquad \leqq \max \, (| 1 - \alpha |, | \beta - 1|) \, \| \mathbf{C}(f - g) \|_\infty \\ &\qquad \leqq \max \, (| 1 - \alpha |, | \beta - 1 |) \, \| \mathbf{C} \|_\infty \cdot \| f - g \|_\infty . \end{aligned}$$

Thus **L** is a contraction when $r < 1$. This proves Theorem II with the exception of the last inequality* which follows directly from the fact that $s_2$ can be written as

$$s_2 = s_{20} + \sum_{j=0}^{\infty} [s_{2(j+1)} - s_{2j}] \tag{12}$$

in which for all $j \geqq 1$

$$\| s_{2(j+1)} - s_{2j} \|_\infty = \| \mathbf{L}s_{2j} - \mathbf{L}s_{2(j-1)} \|_\infty \leqq r \| s_{2j} - s_{2(j-1)} \|_\infty .$$

5.2 *Proof of Theorem III:*

With obvious modifications the proof of Theorem II suffices.

## VI. A CASCADE GRAPH FOR APPROXIMATELY DETERMINING THE SIGNAL $s_2$ IN FIG. 3

Suppose that the input signal $s_1$ in Fig. 3 is an element of $\mathcal{K}(\Sigma)$. Then under the assumptions stated in Theorem II, the output signal $s_2$ is an element of $\mathcal{K}(\Sigma)$ and is given by (12) where $s_{20}$ is an arbitrary element of $\mathcal{K}(\Sigma)$. The key inequality that must be satisfied is† (using the first part of Theorem I and assuming that $c, h \in \mathfrak{D}$)

$$\int_{-\infty}^{\infty} | c(t) | \, dt \left(1 + \int_{-\infty}^{\infty} | h(t) | \, dt\right) \cdot \max (| 1 - \alpha |, | \beta - 1 |) = r < 1. \tag{13}$$

If we take $s_{20} = (\mathbf{I} - \mathbf{C})^{-1}s_1$, the sum $\sum_{j=0}^{\infty} [s_{2(j+1)} - s_{2j}]$ represents the nonlinear distortion present in the output. The first term in this series is

$$(s_{21} - s_{20}) = (\mathbf{I} - \mathbf{C})^{-1}\tilde{\varphi}[(\mathbf{I} - \mathbf{C})^{-1}s_1]$$

and, using the inequality in Theorem II, a bound on the error incurred in ignoring the remainder of the series is given by

$$\left\| (s_{21} - s_{20}) - \sum_{j=0}^{\infty} [s_{2(j+1)} - s_{2j}] \right\|_\infty = \| s_{21} - s_2 \|_\infty$$

$$\leqq \frac{r}{1 - r} \| s_{21} - s_{20} \|_\infty .$$

---

* Note that when $s_{20} = 0$, this inequality implies that

$$\| s_2 \|_\infty \leqq (1 - r)^{-1} \| s_{21} \|_\infty .$$

† In physical systems both integrands vanish for negative arguments.

That is, the error in ignoring the remainder of the series is at most $r(1 - r)^{-1}$ times the norm of the first term. Thus if $r$ is sufficiently small the function $(s_{21} - s_{20})$ is a good approximation to the distortion component of the signal $s_2$. Fig. 4 shows the corresponding flow graph for determining $s_{21}$, the approximation to the output signal in Fig. 3. Theorem III leads to analogous results and the same flow graph for the case in which signals belong to $\mathfrak{L}_{2R}$. The essential difference is that in the $\mathfrak{L}_{2R}$ case attention is focused on the energies of the signals.

### 6.1 *Relation of the Graph in Fig. 4 to a Well-Known Engineering Technique*

The flow graph in Fig. 4 characterizes the essence of a well-known engineering technique[5,6,7] for approximately determining the effect of feedback on nonlinear distortion introduced in one stage of an amplifier, when the distortion is "small." In particular, observe that if, as indicated in Fig. 4, $u(t)$ denotes the distortion* produced by the open-loop system

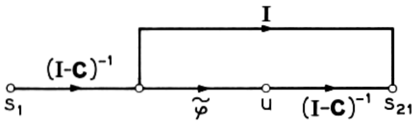


Fig. 4 — Cascade flow graph for approximately determining the output signal in Fig. 3 ($s_{21}$ is the approximation to $s_2$).

with the same "small-signal transmission" and the same output stage as the feedback system in Fig. 3, then the output distortion in Fig. 3 is approximately $(\mathbf{I} - \mathbf{C})^{-1}u$. In engineering terms, feedback is said[5,6,7] to reduce the nonlinear distortion by the amount of the "return difference" $[1 - C(\omega)]$ [i.e., by the formal frequency domain representation of the operator $(\mathbf{I} - \mathbf{C})$].

### 6.2 *Comparison with Desoer's Results*

In an interesting paper[1] Desoer has considered the range of validity of the graph in Fig. 4 for the case in which $\varphi(v,t) = v + \epsilon v^m$, where $\epsilon$ and $m$ are real constants with $m$ an odd positive integer. In his discussion all signals belong to $\mathcal{K}(\Sigma)$ with $\Sigma$ the null set. He considers the analysis of a feedback system of the type shown in Fig. 2(a) and argues that if the norm of the input to the amplifier is sufficiently small and if $|\epsilon|$ is sufficiently small,† then the distortion component of the output

* The "distortion generator" referred to in the usual engineering arguments produces the signal $u(t)$.

† This writer feels that some additional restriction on $|\epsilon|$ is necessary in order that Desoer's condition ($B$) be satisfied. It would suffice to assume that $|\epsilon| < 3/16$.

signal is given by the sum of an infinite series in which the norm of each term is less than one-fourth the norm of the preceding term. When the system is characterized by a flow graph of the type shown in Fig. 3 [as indicated earlier, the analysis of the seemingly more complicated situation in Fig. 2(a) can be reduced at once to a consideration of this type of graph], the first term in the series is the distortion determined from Fig. 4.

Desoer focuses attention on simplifications that can be exploited in cases of engineering interest. According to him it is a matter of experience that for a typical low-pass feedback amplifier $\| \mathbf{C} \|_\infty = -C(0)$ [observe that in general $\| \mathbf{C} \|_\infty \geqq \sup_\omega | C(\omega) |$]. In addition he presents a heuristic argument to support the claim that in such amplifiers $\| (\mathbf{I} - \mathbf{C})^{-1} \|_\infty$ is approximately equal to 2. With

$$1 + \int_{-\infty}^{\infty} | h(t) | \, dt = \| (\mathbf{I} - \mathbf{C})^{-1} \|_\infty = 2,$$

the condition that $r$ [in (13)] be less than ¼ (i.e., the condition corresponding to Desoer's criterion for determining the applicability of the graph in Fig. 4) is

$$\max (| 1 - \alpha |, | \beta - 1 |) < \frac{1}{8 \| \mathbf{C} \|_\infty}. \tag{14}$$

It is a routine matter to show that in the high loop-gain case (i.e., the case of principal engineering interest) (14) is a much less stringent condition on the permitted degree of nonlinearity than that implied by Desoer's upper bound on $| \epsilon |$ and his input norm bound. For a loop gain of 100 [i.e., $C(0) = -100$] and $m = 3$, the bound on $| \epsilon |$ is such that (14) permits any deviation from unity of the slope of $\varphi(v)$ [i.e., $\max (| 1 - \alpha |, | \beta - 1 |)$] which does not exceed 2,500 times the permitted maximum deviation from unity of the slope of $v + \epsilon v^3$ over the operating range*† implied by Desoer's input norm bound.

### 6.3 *An Extension of Theorem II*

Note that when the loop gain is large the permitted amount of nonlinearity in (14) is quite small. Although it is not clear whether the range of validity of the graph in Fig. 4 can be substantially increased, it is a simple matter to show that the iterates $s_{2n}$ defined in Theorem II

* In the notation of Ref. 1, the signal input to the nonlinear element is $z - \boldsymbol{\mu\beta}\zeta$ and $\| \zeta \| \leqq \frac{4}{3} \| \zeta_1 \| \leqq \frac{8}{3} | \epsilon |$.

† Desoer's bound on $| \epsilon |$ can be considerably improved in the large loop-gain case by assuming that $\zeta$ lies within a ball of much smaller radius (such as $\| \boldsymbol{\mu\beta} \|^{-1}$).

converge to the unique $s_2 \in \mathcal{K}(\Sigma)$ that satisfies $s_2 = \varphi[\mathbf{C}s_2 + s_1]$ even if $r \geqq 1$, provided that

$$\| \mathbf{C}(\mathbf{I} - \mathbf{C})^{-1} \|_\infty \max (| 1 - \alpha |, | \beta - 1|) < 1.$$

This follows from the contraction-mapping fixed-point theorem and the fact that the relation between $s_1$ and $s_2$ can be written as $s_3 = s_1 + \tilde{\varphi}[\mathbf{C}(\mathbf{I} - \mathbf{C})^{-1}s_3 + s_1]$ where $s_3 = (\mathbf{I} - \mathbf{C})s_2$.

In the next section we consider an alternative, but closely related, cascade graph for determining the output distortion in Fig. 3. For a given loop gain (assuming it is large) the alternative graph is valid for much larger amounts of nonlinearity than that indicated in (13) or (14).

## VII. ADDITIONAL RESULTS RELATING TO THE EQUATION $s_2 = \varphi[\mathbf{C}s_2 + s_1]$

In this section we assume that there exists a function $\psi$ such that $\psi[\varphi(x)] = x$ for all real $x$ and $t$. Hence $\psi[s_2] = \mathbf{C}s_2 + s_1$. Specifically

*Definition: Let $\psi(x,t) = x + \tilde{\psi}(x,t)$ be a real-valued function of the real variables $x$ and $t$ such that $\psi(0,t) = 0$ for all $t$, and that for all $t$ and all $x \geqq y$*

$$\gamma(x - y) \leqq \psi(x,t) - \psi(y,t) \leqq \sigma(x - y)$$

*where $\gamma$ and $\sigma$ are real constants.*

*Theorem IV: Let $\psi[f]$ be continuous with respect to $t$ on the complement of $\Sigma$ whenever $f \in \mathcal{K}(\Sigma)$. Let $c(t) \in \mathcal{L}_{1R}$, $C(\omega) \neq 1$, and suppose that*

$$\| (\mathbf{I} - \mathbf{C})^{-1} \|_\infty \max (| 1 - \gamma |, | \sigma - 1 |) = q < 1.$$

*Then for any $s_1 \in \mathcal{K}(\Sigma)$, there exists a unique $s_2 \in \mathcal{K}(\Sigma)$ such that $\psi[s_2] = \mathbf{C}s_2 + s_1$. In fact, $s_2 = \lim_{n\to\infty} \bar{s}_{2n}$ where*

$$\bar{s}_{2n} = -(\mathbf{I} - \mathbf{C})^{-1}\tilde{\psi}[\bar{s}_{2(n-1)}] + (\mathbf{I} - \mathbf{C})^{-1}s_1$$

*and $\bar{s}_{20}$ is an arbitrary element of $\mathcal{K}(\Sigma)$. The nth approximation $\bar{s}_{2n}$ satisfies*

$$\| s_2 - \bar{s}_{2n} \|_\infty \leqq \frac{q^n}{1 - q} \| \bar{s}_{21} - \bar{s}_{20} \|_\infty .$$

The analogous result for the space $\mathcal{L}_{2R}$ is

*Theorem V: Let $c(t) \in \mathcal{L}_{2R}$, $\sup_\omega |C(\omega)| < \infty$, $\inf_\omega |1 - C(\omega)| > 0$, and suppose that*

$$\|(\mathbf{I} - \mathbf{C})^{-1}\|_2 \max (| 1 - \gamma |, | \sigma - 1 |) = q < 1.$$

*Then the conclusion of Theorem IV follows with $\mathcal{K}(\Sigma)$ replaced with $\mathcal{L}_{2R}$ and with the $\mathcal{K}(\Sigma)$ norm replaced with the $\mathcal{L}_{2R}$ norm.*

### 7.1 *Proof of Theorem IV*

In view of the first part of Theorem I and the assumption that $C(\omega) \neq 1$, the equation $\psi[s_2] = \mathbf{C}s_2 + s_1$ can be written as $s_2 = \mathbf{M}s_2$, where $\mathbf{M}$ is the mapping of $\mathcal{K}(\Sigma)$ into itself defined by

$$\mathbf{M}s_2 = -(\mathbf{I} - \mathbf{C})^{-1}\tilde{\psi}[s_2] + (\mathbf{I} - \mathbf{C})^{-1}s_1 .$$

Thus, Theorem IV follows from the contraction-mapping fixed-point theorem and the readily verified fact that

$$\| \mathbf{M}f - \mathbf{M}g \|_\infty \leqq q \| f - g \|_\infty , \qquad q < 1$$

for any $f,g \in \mathcal{K}(\Sigma)$. With obvious modifications this argument suffices to establish Theorem V.

### 7.2 *A Cascade Graph for Approximately Determining the Signal $s_2$ in Fig. 3 when $\| \mathbf{C} \|_\infty$ Is Large*

As in the discussion of Theorem II, if we set $\bar{s}_{20} = (\mathbf{I} - \mathbf{C})^{-1}s_1$ in Theorem IV,

$$s_2 = \bar{s}_{20} + \sum_{j=0}^{\infty} [\bar{s}_{2(j+1)} - \bar{s}_{2j}],$$

in which $\sum_{j=0}^{\infty} [\bar{s}_{2(j+1)} - \bar{s}_{2j}]$ represents the nonlinear distortion component of the output signal in Fig. 3. The first term in this series is

$$(\bar{s}_{21} - \bar{s}_{20}) = -(\mathbf{I} - \mathbf{C})^{-1}\tilde{\psi}[(\mathbf{I} - \mathbf{C})^{-1}s_1], \tag{15}$$

and the error incurred in ignoring the remainder of the series $\left(\text{i.e.}, \left\| (\bar{s}_{21} - \bar{s}_{20}) - \sum_{j=0}^{\infty} [\bar{s}_{2(j+1)} - \bar{s}_{2j}] \right\|_\infty \right)$ is, as in Section VI, at most $q(1 - q)^{-1}$ times the norm of the first term [(i.e., of (15)].

Thus if

$$q = \| (\mathbf{I} - \mathbf{C})^{-1} \|_\infty \max (| 1 - \gamma |, | \sigma - 1 |) \tag{16}$$

is sufficiently small, $(\bar{s}_{21} - \bar{s}_{20})$ is a good approximation to the distortion component of $s_2$. Fig. 5 shows the corresponding flow graph* for determining $\bar{s}_{21}$, the approximation to the output signal $s_2$ in Fig. 3.

* An analogous interpretation of Theorem V leads to the same flow graph.

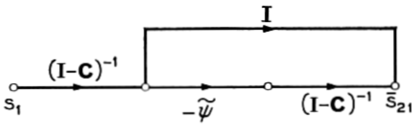


Fig. 5 — Alternative cascade flow graph for approximately determining the output signal in Fig. 3 ($\bar{s}_{21}$ is the approximation to $s_2$).

Observe that this graph differs* from the one in Fig. 4 only in that $\tilde{\varphi}$ is replaced with $-\tilde{\psi}$. However the expression for $q$ in (16), unlike the corresponding expression for $r$ in Theorem II, does not contain the factor $\| \mathbf{C} \|_\infty$, which is a measure of the amount of feedback present in the system (as indicated earlier, in low-pass feedback amplifiers typically $\| \mathbf{C} \|_\infty = | C(0) |$).

The condition that $q$ in (16) be less than $\frac{1}{4}$ when $\| (\mathbf{I} - \mathbf{C})^{-1} \|_\infty = 2$ (i.e., the condition corresponding to Desoer's practical criterion for determining the applicability of the graph in Fig. 4) is clearly†

$$\max(|1 - \gamma|, |\sigma - 1|) < 1/8.$$

## VIII. ACKNOWLEDGMENT

The writer is indebted to V. E. Beneš for reading the draft.

## APPENDIX

The following inequality, in which $\mathfrak{N}$ denotes either $\mathcal{K}(\Sigma)$ or $\mathcal{L}_{2R}$, can be used in some cases to compare the output signals in Figs. 4 and 5

$$\| \mathbf{Q}\tilde{\varphi}[f] + \mathbf{Q}\tilde{\psi}[f] \| \leqq \frac{p}{1 - p} \| \mathbf{Q}\tilde{\psi}[f] \| \tag{17}$$

where $f$ is an arbitrary element of $\mathfrak{N}$, $\| \cdot \|$ denotes the norm for the space $\mathfrak{N}$, and $\mathbf{Q}$ is any linear operator defined on $\mathfrak{N}$ such that

$$\| \mathbf{Q} \| \cdot \| \mathbf{Q}^{-1} \| \max(|1 - \gamma|, |\sigma - 1|) = p < 1.$$

---

* In the Appendix a general inequality is presented which in some cases can be used to bound the norm of the difference between the two approximations to the distortion component of $s_2$, $(s_{21} - s_{20})$ in Fig. 4 and $(\bar{s}_{21} - \bar{s}_{20})$ in Fig. 5, in terms of the norm of $(\bar{s}_{21} - \bar{s}_{20})$. The inequality is not applicable when $\| \mathbf{C} \|_\infty$ is large unless $\max(|1 - \gamma|, |\sigma - 1|)$ is sufficiently small.

† It is sometimes desirable to consider the unrealistic and very much simpler situation in Fig. 3 when $\mathbf{C}$ represents multiplication by a real constant $c$. In that case the relation between $s_1$ and $s_2$ can be written as $s_2 = c^{-1}\psi[s_2] - c^{-1}s_1$, from which it is evident that, provided $|c|$ is sufficiently large, the contraction-mapping fixed-point theorem is applicable with a small contraction constant even when $\psi$ represents a highly nonlinear element.

Inequality (17) follows from the fact that $g = \mathbf{Q}\varphi[f]$ is the fixed point of the contraction-mapping $\mathbf{N}$ (with contraction constant $p$) defined by $\mathbf{N}g = \mathbf{Q}f - \mathbf{Q}\tilde{\psi}[\mathbf{Q}^{-1}g]$.


REFERENCES

1. Desoer, C. A., Nonlinear Distortion in Feedback Amplifiers, I.R.E. Trans. on Circuit Theory, **CT-9,** March, 1962, pp. 2–6.
2. Kolmogorov, A. N. and Fomin, S. V., *Elements of the Theory of Functions and Functional Analysis*, New York, Graylock Press, 1957.
3. Sandberg, I. W., On The Properties of Some Systems That Distort Signals — II, to be published.
4. Pitt, H. R., *Tauberian Theorems*, Bombay, Oxford University Press, 1958, p. 99.
5. Bode, H. W., *Network Analysis and Feedback Amplifier Design*, New York, D. Van Nostrand, 1945.
6. Gray, T. S., *Applied Electronics*, New York, John Wiley and Sons, 1954.
7. Angelo, E. J., Jr., *Electronic Circuits*, New York, McGraw-Hill, 1958.