SEL-67-016

THE OPERATIONAL CALCULUS OF JAN MIKUSIŃSKI

by

R. W. Newcomb

February 1967

Technical Report No. 6560-11

This work was prepared
under the Office of Naval Research
Contract Nonr 225(83)

Systems Theory Laboratory
Stanford Electronics Laboratories
Stanford University Stanford, California

THE OPERATIONAL CALCULUS OF JAN MIKUSIŃSKI

by

R. W. Newcomb

CONTENTS

I. Introduction, notation for functions

II. The Sets $\mathcal{C}$ and $\mathcal{L}$, The Ring $\mathcal{R}_c$

- a) $\mathcal{C}$ = Continuous Functions
- $\mathcal{L}$ = Lebesgue Integrable Functions
- c) Product in $\mathcal{C}$ as Convolution to Give a Ring
- d) Appendix - Rings, Integral Domains

III. Operators

- a) Extension of $\mathcal{R}_c$ to Field $\mathcal{O}$ of Operators
- b) Elements in $\mathcal{O}$

IV. Convergence in $\mathcal{O}$, The Impulse "Function" and Derivatives

V. Integrodifferential Equations

- a) Eq. 5.2), Forcing Function in $\mathcal{L}$
- b) Definition of Solution
- c) Existence of Solution of Eq. 5.2) in $\mathcal{L}$
- d) Impulse Response

VI. Extensions and Comments

- a) Time-variable, Distributional, and n-Variable Cases
- b) "Disjointness" of Operators and Distributions

VII. Critique

- a) Comparison with Laplace transform
- b) Comparison with Distributions

VIII. Acknowledgments

IX. References

ABSTRACT

To operate within ones domain
Is a function that's least understood
As the calculus herein contained.
Abstracted to the very most prime details
Is the theory Mikusiński propounds
With comments and more recent extends
In the end does the domain explain.

PROLOGUE

Who, if I shouted, among the hierarchy of angels
would hear me? And supposing one of them
took me suddenly to his heart, I would perish
before his stronger existence.

Rilke, The First Elegy [RI 1, p.3]

I. INTRODUCTION

The operational calculus, as developed by Mikusiński [MI 1], is designed to solve more general differential equations than the Laplace transform method. At the same time it gives a rigorous definition of the impulse function and all of its derivatives.

The calculus is quite similar to that of Heaviside but is also rigorous. It is developed in the following manner. We begin by considering a restricted class of continuous functions $\mathcal{C}$. The set of functions $\mathcal{C}$ is then made into a ring $\mathfrak{R}_c$ by defining multiplication as convolution and using addition in the standard manner. $\mathfrak{R}_c$ turns out to be a commutative ring with no divisors of zero and no unit element. It is then extended to its field of quotients to yield the field $\mathcal{O}$ of operators. $\mathcal{O}$ contains a unit element, the impulse function, as well as many other entities which are not functions. The members of $\mathcal{O}$ can then be used to solve constant coefficient integrodifferential equations with the most general types of forcing functions.

In the following we will generally adhere to Mikusiński's notation which has the following convention concerning functions.

$$\begin{aligned} f &= \{f(t)\} = \text{the function } f \\ f(t) &= \text{the function } f \text{ evaluated at the point } t \end{aligned} \tag{1.1}$$

One of the main reasons for making this distinction between $\{f(t)\}$ and f(t) is to avoid trouble with constant functions. Thus consider f = {1} which is graphed in Fig. 1.1a). Comparing this with Fig. 1.1b) shows that f = {1} and the number 1 are two distinct concepts; one can be thought of as a line, and the other as a point.

a)
f = {1}

b)
The Number 1

Figure 1.1

Comparison of f = {1} and the Number 1

II. THE SETS $\mathcal{C}$ AND $\mathcal{L}$, THE RING $\mathfrak{R}_c$

Let R denote the set of real numbers and let R_T denote the set of nonnegative real numbers less than the positive number T (that is $t \epsilon R_T$ if and only if $0 \leq t < T \leq \infty$). Further, in the following we will assume that the letter f, possibly with subscripts, will denote a complex-valued function defined over R. We will base our calculus on the set $\mathcal{C}$ of functions continuous on R_∞ [MI 1, p. 2].

Definition 2.1: $f \epsilon \mathcal{C}$ if

1) $f(t)$ is defined for each $t \epsilon R$
2) f is continuous on R_∞
3) $f(t) = 0$ for $t \epsilon C(R_\infty)$

Here $C(\ )$ denotes the complement with respect to R. Note that f need not be continuous at $t = 0$ when it is considered on R.

When considering differential equations we will be interested in discontinuous functions. For this we will use the set $\mathcal{L}$ of locally integrable functions [MI 1, p. 345].

Definition 2.2: $f \epsilon \mathcal{L}$ if

1) f is Lebesgue integrable in R_T for every finite $T > 0$
2) $f(t) = 0$ for $t \epsilon C(R_\infty)$
3) $f_1 = f_2$ if f_1 and f_2 differ only on a set of measure zero.

The need for the identification in 3) will become clear after $\mathcal{O}$ is defined, but it is standard in integration theory. Note that $\mathcal{C} \subset \mathcal{L}$ as every $f \epsilon \mathcal{C}$ is Lebesgue integrable in R_T and if $f_1 \epsilon \mathcal{C}$ and f_2 differs by at most one point from f_1 then $f_2 \not\epsilon \mathcal{C}$.

Consider the following examples for which the f are all assumed to be zero for $t \epsilon C(R_\infty)$.

Example 2.1: Let $f = \{a\}$ for $t \epsilon R_\infty$, a = complex constant; $f \epsilon \mathcal{C} \subset \mathcal{L}$.

Example 2.2: Let $f = \{e^{-at}\}$ for $t \epsilon R_\infty$, a = complex constant; $f \epsilon \mathcal{C} \subset \mathcal{L}$.

Example 2.3: Let $f = \{t^{-1/2}\}$ for $t \epsilon R_\infty$, $f \epsilon \mathcal{L}$, $f \not\epsilon \mathcal{C}$

Example 2.4: $f = \left\{\begin{matrix} 0 & t = \text{rational} \\ 1 & t = \text{irrational} \end{matrix}\right\}$ for $t \epsilon R_\infty$; $f \epsilon \mathcal{L}$, $f \not\epsilon \mathcal{C}$

Note that f is identified with $\{1\}$ in $\mathcal{L}$

Example 2.5: $f = \{t^{-1}\}$ for $t \epsilon R_\infty$; $f \not\epsilon \mathcal{L}$, $f \not\epsilon \mathcal{C}$.

If $f \in \mathfrak{L}$ it will be referred to by its functional values for $t \in R_\infty$, unless some confusion can arise as to the values in $C(R_\infty)$. Thus the functions of Examples 2.1, 2.2, 2.3 will be denoted by $\{a\}$, $\{e^{-at}\}$, $\{t^{-1/2}\}$ respectively; a more meaningful notation would incorporate the unit step function, H_0 in Mikusiński's notation, but this is rather cumbersome for our purposes.

We now introduce the operations of addition (in the standard way), and multiplication (as convolution), into $\mathcal{C}$.

<u>Definition</u> 2.3: For $f_1 \in \mathcal{C}$ and $f_2 \in \mathcal{C}$ define

1) $f_1+f_2 = \{f_1(t)+f_2(t)\}$

2) $f_1 f_2 = f_1 \cdot f_2 = f_1 * f_2 = \{\int_0^t f_1(t-\tau)f_2(\tau)d\tau\}$

The resulting system will be denoted by $\mathcal{R}_c$. Here and in the following we will understand all integrals to be Lebesque integrals. We can justify this type of multiplication by noting that we wish to consider linear systems where the convolution of the excitation with the impulse response plays a key role in the resolution. Also we could equally well have given a definition like 2.3 for $\mathfrak{L}$. However, this yields the same operational calculus, or field $\mathcal{O}$, while merely complicating matters. The following fact, called Titchmarsh's Theorem, is indispensable.

<u>Lemma</u> 2.1: If $f_1 \in \mathfrak{L}$, $f_2 \in \mathfrak{L}$ and if $\{\int_0^t f_1(t-\tau)f_2(\tau)d\tau\} = \{0\}$ almost everywhere in R_∞ then at least one of f_1 or f_2 vanishes almost everywhere.

The proof is quite complicated but is found in [MI 1, pp. 15-23 and 346]. Note that the almost everywhere can actually be deleted from Lemma 2.1 since we have agreed to call f_1 zero if it vanishes almost everywhere. Also note that if we allow functions which are nonzero in $C(R_\infty)$ and take the limits of integration as $-\infty$, ∞ for the convolution, then Lemma 2.1 does not hold. This is seen by letting $f_1 = \{1\}$, $f_2 = \{t/(1+t^2)\}$ for all $t \in R$, then

$$\{\lim_{\alpha \to \infty} \int_{-\alpha}^{\alpha} [\tau/(1+\tau^2)]d\tau\} = \{0\} \qquad (2.1)$$

(note that $\int_{-\infty}^{\infty} [\tau/(1+\tau^2)]d\tau$ does not exist as a true Lebesque integral).

As a result of the meaning of multiplication and Titchmarsh's Theorem, the system is an integral domain, see the appendix to this section (the integral domain is as defined by [MC 1, p. 15], not as defined by [BI 1, p. 1] where a unit is needed).

Theorem 2.1: $\mathfrak{R}_c$ has the following properties:

$$\left.\begin{array}{ll} 1) & \text{It is a commutative ring} \\ 2) & \text{It has no proper divisors of zero} \\ 3) & \text{It has no unit element} \end{array}\right\} = \text{integral domain}$$

Proof: That $\mathfrak{R}_c$ is a ring is seen in the following manner. The sum and convolution of two functions in $\mathcal{C}$ is again in $\mathcal{C}$. Clearly addition is commutative and associative, and each function in $\mathcal{C}$ has an additive inverse. We can show that $(f_1f_2)f_3 = f_1(f_2f_3)$ by changing the order of integration. Thus (see Fig. 2.1 for limit changes)

$$(f_1f_2)f_3 = \{\int_0^t [\int_0^{t-\tau} f_1(t-\tau-\sigma)f_2(\sigma)d\sigma]f_3(\tau)d\tau\}, \text{ let } \sigma = \omega-\tau \quad (2.2)$$

$$= \{\int_0^t [\int_\tau^t f_1(t-\omega)f_2(\omega-\tau)d\omega]f_3(\tau)d\tau\}$$

$$= \{\iint_S f_1(t-\omega)f_2(\omega-\tau)f_3(\tau)dA\}$$

$$= \{\int_0^t [\int_0^\omega f_2(\omega-\tau)f_3(\tau)d\tau]f_1(t-\omega)d\omega\}$$

$$= f_1(f_2f_3)$$

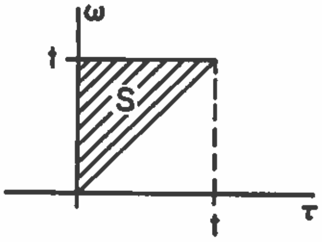

Figure 2.1

Integration Area

It is also clear that $f_1(f_2+f_3) = f_1f_2+f_1f_3$ and $(f_1+f_2)f_3 = f_1f_3+f_2f_3$. Thus $\mathfrak{R}_c$ is a ring. It is commutative because f_1f_2 is seen to equal f_2f_1 by inspection of the convolution integral. Lemma 2.1 shows that $\mathfrak{R}_c$ has no divisors of zero. To see that $\mathfrak{R}_c$ has no unit element we can assume the contrary, that is, assume $e\epsilon\mathfrak{R}_c$ to be a unit element. Then

$$e\{1\} = \{\int_0^t e(\tau)d\tau\}$$

which can not equal $\{1\}$ for any continuous $\{e(t)\}$ as is seen by considering $t = 0$. Q.E.D.

II-a. APPENDIX TO II

Here we define the algebraic concepts associated with integral domains [MC 1].

Definition $a_{II}.1$: A set $\mathfrak{R}$ of elements $r_1, r_2, \ldots,$ is called a ring if two binary operations $+$ and $\cdot$ are defined and conform to the following laws:

1) $r_1+r_2 \epsilon \mathfrak{R}$, $r_1 \cdot r_2 \epsilon \mathfrak{R}$ for every r_1 and $r_2 \epsilon \mathfrak{R}$
2) $r_1+(r_2+r_3) = (r_1+r_2)+r_3$
3) $r_1+r_2 = r_2+r_1$
4) $r_1+x = r_2$ has a solution $x \epsilon \mathfrak{R}$
5) $r_1 \cdot (r_2 \cdot r_3) = (r_1 \cdot r_2) \cdot r_3$
6) $r_1 \cdot (r_2+r_3) = r_1 \cdot r_2 + r_1 \cdot r_3$; $(r_1+r_2) \cdot r_3 = r_1 \cdot r_3 + r_2 \cdot r_3$

A ring need not have a unit element, that is an e such that $e \cdot r_1 = r_1 \cdot e = r_1$ for all $r_1 \epsilon \mathfrak{R}$. A ring is called a commutative ring if $r_1 \cdot r_2 = r_2 \cdot r_1$ for all r_1 and $r_2 \epsilon \mathfrak{R}$. By property 4) there is an element 0 such that $r_1+0 = r_1$ for all $r_1 \epsilon \mathfrak{R}$. A nonzero element r_1 such that there exists a nonzero element r_o such that either $r_o \cdot r_1 = 0$ or $r_1 \cdot r_o = 0$ is said to be a proper divisor of zero.

Definition $a_{II}.2$: A commutative ring with no proper divisors of zero is called an integral domain.

The following examples illustrate these concepts.

Example $a_{II}.1$: Consider the set of integers with the usual laws of addition and multiplication. This is a commutative ring with a unit and no proper divisors of zero. The same is true of the rational numbers, the real numbers R, and the complex numbers.

Example $a_{II}.2$: Consider the set of 2×2 matrices with elements in R (or any other ring with a unit). This is a noncommutative ring with a unit, and it has proper divisors of zero. Thus

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \neq \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \text{ (noncommutivity)}$$

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{(zero divisor)}$$

Example a_{II}.3: The set of all distributions over R does not form a ring with convolution as $\cdot$ since associativity does not hold [SC 1, vol. 2, p. 21]

$$(\{1\}*\delta^{(1)})*H_o = 0, \quad \{1\}*(\delta^{(1)}*H_o) = \{1\}*\delta = \{1\}$$

Here δ is the impulse, H_o is the unit step function, and $\{1\}$ is defined and constant for $-\infty \leq t \leq \infty$; note the presence of a zero divisor (compare Eq. 2.1)).

III. OPERATORS

Since no element in $\mathfrak{R}_c$ has no unit element, there is no element in $\mathfrak{R}_c$ which has a multiplicative inverse. However $\mathfrak{R}_c$, being an integral domain, can be inbedded in a field of quotients [MC 1, p. 94]. The elements of this field are Mikusiński's operators [MI 1, p. 25].

Definition 3.1: The field whose elements are f_1/f_2 with f_1 and $f_2 \epsilon \mathfrak{R}_c$, $f_2 \neq \{0\}$ is denoted by $\mathcal{O}$. Its elements are called operators.

$\mathcal{O}$ is then obtained in the same manner that the field of rationals is obtained from the ring of integers. The construction considers two functions f_1 and $f_2 \neq \{0\}$ as pairs (f_1,f_2) written f_1/f_2 where the various operations $(=, +, \cdot)$ on these operators are defined through the following equations.

Properties 3.1: For every $f_1, f_2, f_3, f_4 \epsilon \mathfrak{R}_c$, $f_2 \neq \{0\}$, $f_4 \neq \{0\}$ we have

1) $f_1/f_2 = f_3/f_4$ if and only if $f_1f_4 = f_2f_3$

2) $(f_1/f_2)(f_3/f_4) = (f_1f_3)/(f_2f_4)$

3) $(f_1/f_2) + (f_3/f_4) = (f_1f_4 + f_2f_3)/(f_2f_4)$

Our position is as follows. We were not always allowed to divide in $\mathfrak{R}_c$ so we have extended $\mathfrak{R}_c$ to $\mathcal{O}$. $\mathcal{O}$ is constructed such that every nonzero element $\sigma \epsilon \mathcal{O}$ has a multiplicative inverse $\sigma^{-1} \epsilon \mathcal{O}$. Thus there is a unit element $1 \epsilon \mathcal{O}$. (Note the difference between the operators $\{1\}$ and 1, the first is a function in $\mathcal{C}$, the latter is not; 1 is as the impulse). Likewise every equation $x\sigma_1 = \sigma_2$ can be solved for $x = \sigma_2\sigma_1^{-1} \epsilon \mathcal{O}$. It now remains to interpret some of the operators and show their usefulness. We will see that the functions of $\mathcal{C}$ and $\mathcal{L}$ are in $\mathcal{O}$, as well as the complex constants. Likewise, there are operators which perform the operations of integration and differentiation.

Property 3.2: If $f \epsilon \mathcal{C}$ then $f \epsilon \mathcal{O}$; that is $\mathfrak{R}_c \subset \mathcal{O}$.

Proof: We first note that $\mathcal{O}$ is actually only defined up to isomorphism. Thus we write $f = (f/f_1)(f_1f_1/f_1)$ for any nonzero $f_1 \epsilon \mathfrak{R}_c$. With this convention f is clearly in $\mathcal{O}$. Q.E.D.

One operator of importance is $\{1\} \epsilon \mathcal{C}$.

Definition 3.2: The operator

$$\ell = \{1\}, \quad \{1\} \epsilon \mathcal{C}$$

is called the <u>integral operator</u>.

This name is clearly motivated by the fact that

$$\ell f = \{\int_0^t f(\tau)d\tau\}, \quad f \epsilon \Re_c \tag{3.1}$$

of course ℓ is identified also with the unit step function H_o; $\ell = H_o = \{1\} \epsilon \mathcal{C}$.

<u>Property</u> 3.3: If $f \epsilon \mathfrak{L}$ then $f \epsilon \mathcal{O}$

Proof: Defining ℓf by Eq. 3.1) we know that $\ell f \epsilon \Re_c \subset \mathcal{O}$ because the indefinite integral of a Lebesgue integrable function is continuous [MC 2, p. 159]. Then $[(\ell f)/\ell] = f$ gives $f \epsilon \mathcal{O}$. Note that $\ell f_1 = \ell f_2$ if f_1 and f_2 differ on a set of measure zero; this explains condition 3) of Definition 2.2. Q.E.D.

<u>Property</u> 3.4: If a is a complex constant then $a \epsilon \mathcal{O}$.

Proof: Consider operators of the form $a = \{\alpha\}/\{1\}$ for $\{\alpha\}, \{1\} \epsilon \Re_c$, α a complex constant. We have $(\{\alpha_1\}/\{1\})+(\{\alpha_2\}/\{1\}) = (\{\alpha_1+{}_2\}/\{1\} = a_1+a_2$; $(\{\alpha_1\}/\{1\})(\{\alpha_2\}/\{1\}) = (\{\alpha_1\alpha_2 t\}/\{1\}^2) = (\{1\}\{\alpha_1\alpha_2\}/\{1\}^2) = a_1a_2$. Thus the operators $a = \{\alpha\}/\{1\}$ behave just like complex numbers. Since the field of quotients is only determined up to isomorphism we can call these operators complex numbers. Note, however, the difference between a_1a_2 and $\{a_1\}\{a_2\} = \{a_1a_2t\}$. Q.E.D.

As indicated by the above proof we then write

$$a = \{a\}/\ell \tag{3.2}$$

as a direct consequence of which we have

$$a\ell = \{a\} = a\{1\} \tag{3.3}$$

In $\mathcal{O}$ every nonzero element has a multiplicative inverse. Of particular importance is the inverse of the integral operator.

<u>Definition</u> 3.3: The operator

$$s = \ell^{-1}$$

is called the differential operator.

The name assigned to s is justified by the following theorem, which results from the fact that an absolutely continuous function possesses a derivative almost everywhere [MC 2, p. 20] [MI 1, p. 349]. Note that we can not extend this to arbitrary $f\in\mathcal{C}$ since a continuous function need not possess a derivative almost everywhere [MC 2, p. 44]. The result is of course that familiar from Laplace transform theory.

Theorem 3.1: If $f\in\mathcal{C}$ is absolutely continuous with f' the derivative then

$$sf = f'+f(0)$$

(here $f(0)$ is a constant $\in\mathcal{O}$, $f\in\mathcal{C}$, $f'\in\mathcal{L}$).

Proof: For an absolutely continuous function we have [MC 1, p. 208]

$$f = \{\int_0^t f'(\tau)d\tau+f(0)\} = \ell f'+f(0)\ell$$

or

$$sf = \ell^{-1}f = f'+f(0)$$

Q.E.D.

Applying this to a function with $n-1$ absolutely continuous derivatives gives [MI 1, p. 349].

Corollary 3.1: If $f\in\mathcal{C}$ has its $(n-1)^{th}$ derivative absolutely continuous then

$$f^{(n)} = s^n f - s^{n-1}f(0) - \dots - f^{(n-1)}(0)$$

For solving differential equations the exponential functions are needed. We have, with $\Gamma(\lambda)$ the Gamma function,

$$(s+a)^{-\lambda} = \{t^{\lambda-1}e^{-at}/\Gamma(\lambda)\} \quad \text{for } \lambda > 0 \qquad (3.4)$$

$$a = \text{complex constant}$$

Proof: We have, by Theorem 3.1, $(s+a)\{e^{-at}\} = \{-ae^{-at}\}+1+\{ae^{-at}\} = 1$ which verifies the case $\lambda = 1$. For λ an integer ≥ 1 the above formula is verified by induction. For nonintegral λ the above is used as the definition of $(s+a)^{-\lambda}$. This definition is a consistent one since exponents then add [MI 1, p. 112]. Q.E.D.

A special case of Eq. 3.4) occurs when $a = 0$. Then we find

$$\ell^{\lambda} = \{t^{\lambda-1}/\Gamma(\lambda)\}; \quad \lambda > 0 \tag{3.5}$$

This then gives rise to a generalization of the integral [CO 1, p. 340] since ℓ^{λ} is valid for other than integer λ.

The Heaviside function (shifted unit step function) is defined by

$$H_{\lambda}(t) = \begin{cases} 0 & -\infty < t < \lambda \\ 1 & 0 \leq \lambda \leq t < \infty \end{cases} \tag{3.6}$$

and we see that $H_{\lambda} \epsilon \mathcal{O}$ since $H_{\lambda} \epsilon \mathcal{L}$. The important operator, which corresponds to the impulse applied at $t = \lambda$, is now defined [MI 1, p. 116].

Definition 3.4: The operator

$$h^{\lambda} = sH_{\lambda} \quad \text{for} \quad \lambda > 0; \; h^{0} = 1$$

is called the translation operator. After a derivative with respect to λ is defined by the standard manner of defining derivatives it is seen that [MI 1, p. 192]

$$h^{\lambda} = e^{-s\lambda} \tag{3.7}$$

The name is justified since, for $f \epsilon \mathcal{L}$

$$h^{\lambda}\{f(t)\} = \begin{cases} 0 & \text{for} \quad 0 \leq t \epsilon \lambda \\ f(t-\lambda) & \text{for} \quad 0 \leq \lambda < t \end{cases} \tag{3.8}$$

Proof: $h^{\lambda}f = sH_{\lambda}f = s\{\int_{0}^{t} f(t-\tau)H_{\lambda}(\tau)d\tau\} = s\{\int_{\lambda}^{t} f(t-\tau)d\tau\} = s\{\int_{0}^{t-\lambda} f(\omega)d\omega\}$

with $\omega = t-\tau$. This is clearly $\{0\}$ for $t < \lambda$ and $\{f(t-\lambda)\}$ for $t > \lambda$. Q.E.D.

Note that $h^{0}\{f(t)\} = \{f(t)\}$ and hence the unit $1 = h^{0}$ is the impulse "function."

Consequently the inverse $h^{-\lambda}$ of h^{λ}, which is in $\mathcal{O}$, translates a function to the left by λ. Considering $h^{-\lambda_2}h^{\lambda_1}f$, where $\lambda_2 > \lambda_1$ and $f \epsilon \mathcal{L}$ we see that there are functions in $\mathcal{O}$ which are nonzero for $-\lambda < t$ for every $\lambda > 0$ as illustrated in Fig. 3.1.

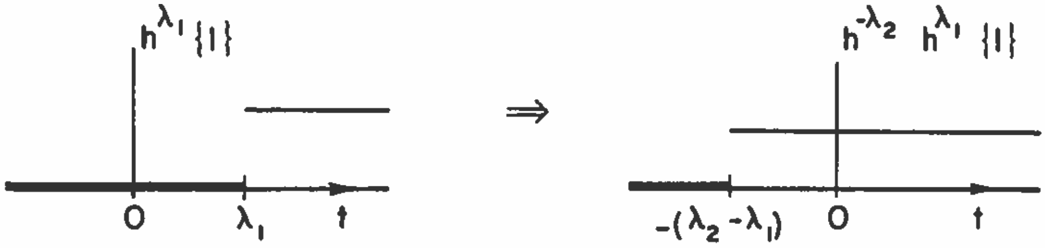

Figure 3.1

Illustration of Type of Functions in $\mathcal{O}$

The difference between these operators and other operators, such as those in the Heaviside calculus, should be noted. Formerly operators were abstract symbols which had to "operate" on something. However, this is not the case in this theory; some operators are numbers, some functions, and others simply operators (such as the "impulse," $1 = h^{0}$). We thus have

$$\text{Complex numbers} \subset \mathcal{C} \subset \mathcal{L} \subset \mathcal{O}$$

Note that all properties of $\mathcal{O}$ are deduced from $\mathcal{C}$ and the introduction of $\mathcal{L}$ is only needed to obtain useful interpretations. Also it is to be noted that s is not a complex number as in the Laplace transform theory; no region of convergence need by considered.

IV. CONVERGENCE IN $\mathcal{O}$, THE IMPULSE "FUNCTION" AND DERIVATIVES

We will show how the operator h^{λ} can be approximated by other operators which are functions. We first need some preliminaries. In order to consider functions defined for negative t, let R_T^{-T} denote the set of real numbers t in $-T < t < T$.

Definition 4.1: A sequence of functions, Lebesgue integrable on R_T^{-T}, is called almost uniformly convergent if it is uniformly convergent in R_T^{-T} for every finite T.

Example 4.1: $f_n(t) = e^{-t/n}$ defines an almost uniformly convergent sequence which is not uniformly convergent in any infinite interval (as $|t|$ can be made greater than any n).

With this we can define convergence in $\mathcal{O}$ [MI 1, p. 144].

Definition 4.2: A sequence of operators defined by σ_n has a limit σ_o if there exists an operator q such that the sequence defined by σ_n/q is almost uniformly convergent to σ_o/q. We write

$$\lim_{n\to\infty} \sigma_n = \sigma_o = q \lim_{n\to\infty} (\sigma_n/q)$$

(note that the convergence on the left is in terms of operators, while that on the right is in terms of functions). This limit is unique, and if the σ_n are almost uniformly convergent functions it agrees with the "functional" limit, since q can be taken as 1 [MI 1, p. 145]. This notion is illustrated by the following examples.

Example 4.2: $\lim_{n\to\infty} \{\cos(nt)\} = 0$ in $\mathcal{O}$

since

$$s \lim_{n\to\infty} \ell\{\cos(nt)\} = s \lim_{n\to\infty} \{(1/n)\int_0^{nt} \cos x dx\} = s \lim_{n\to\infty} \{[\sin(nt)]/n\}$$

Note that cos(nt) has no functional limit.

Example 4.3: $\lim_{n\to\infty} h^n = 0$ in $\mathcal{O}$

since

$$s \lim_{n\to\infty} \ell H_n = s \lim_{n\to\infty} \left\{ \begin{matrix} 0 \text{ if } t < n \\ t-n \text{ if } t > n \end{matrix} \right\} = 0 \quad \text{for } t \text{ bounded}$$

Now h^λ/s^2 is obtained from $s^{-2} = \{t\} \epsilon \mathcal{C}$ by translating by λ. Thus consider

$$f_n(\lambda) = s^{-2}[\tfrac{n}{2}(h^{\lambda-1/n} - h^{\lambda+1/n})] \qquad (4.1)$$

This converges almost uniformly to H_λ as is seen by Fig. 4.1. (Remember our convention removes the difference at $t = \lambda$).

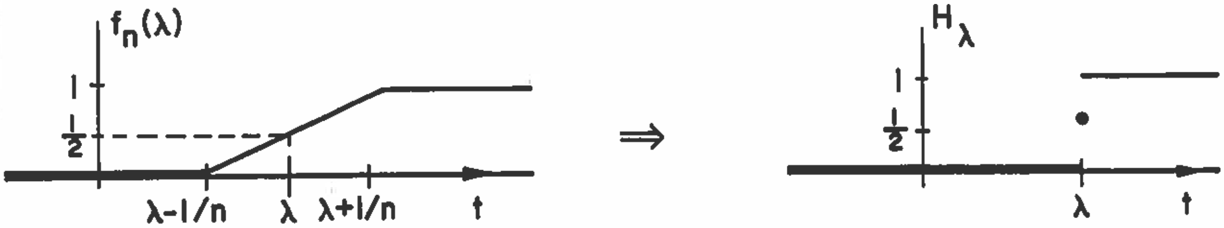

Figure 4.1

Almost Uniform Convergence of $f_n(\lambda)$

We then have

$$\lim_{n\to\infty} sf_n(\lambda) = s \lim_{n\to\infty} f_n(\lambda) = sH_\lambda = h^\lambda \qquad (4.2)$$

or

$$\lim_{n\to\infty} s^{-1}[\tfrac{n}{2}(h^{\lambda-1/n} - h^{\lambda+1/n})] = h^\lambda \qquad (4.3)$$

This is represented graphically in Fig. 4.2.

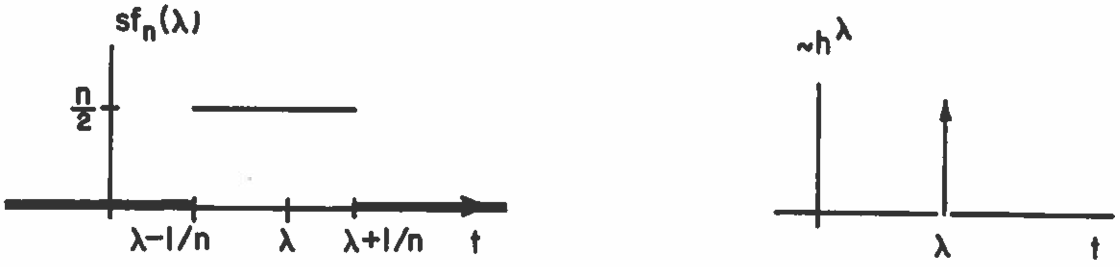

Figure 4.2

Convergence to h^λ

Consequently h^λ can be considered an impulse and can be operationally approached by step functions [MI 1, p. 122]. Of course many other functions can be made to operationally approach h^λ, and any order derivative of the impulse can be considered by $s^m h^\lambda$.

V. INTEGRODIFFERENTIAL EQUATIONS

We can apply the preceeding material to the solution of linear integrodifferential equations with constant coefficients. Let

$$x^{(j)} = d^j[x(t)]/dt^j \quad \text{for integer } j > 0 \qquad (5.1)$$

$$x^{(-j)} = \ell^j x \quad \text{for integer } j > 0$$

$$x^{(0)} = x$$

We then wish to solve equations of the form

$$\sum_{i=-m}^{n} a_i x^{(i)} = f; \quad f \in \mathcal{L}, \; n \geq -m \qquad (5.2)$$

where all a_i are real constants, $a_{-m} \neq 0$, $a_n \neq 0$ (however m or n may be negative). If $n > 0$ then Eq. 5.2) is assumed to be subject to the initial conditions

$$x(0) = \gamma_o, \ldots, x^{(n-1)}(0) = \gamma_{n-1}; \quad n > 0 \qquad (5.3)$$

[note that (initial) integration constants (as functions) are contained in f and that derivatives at zero are evaluated by approaching from the right].

We must state exactly what we mean by a solution of Eq. 5.2) [MI 1, p. 115].

Definition 5.1: An operator x is called a solution of Eq. 5.2) if

1) $x^{(n-1)}$ is absolutely continuous in R_∞

2) Eq. 5.2) is satisfied almost everywhere in R_∞

3) Eq. 5.3) is satisfied for $n > 0$.

Example 5.1: $x = sf$ is a solution of $x^{(-1)} = f \in \mathcal{L}$ since, with $n = -1$, $x^{(-2)} = sf/s^2 = f/s$ which is absolutely continuous as the integral of $f \in \mathcal{L}$. Note that $x = sf \notin \mathcal{L}$ in general.

Theorem 5.1: A unique solution of Eq. 5.2) exists in $\mathcal{O}$ and is given by (if $n \leq 0$ the sum in the numerator is absent)

$$x = [f + \sum_{i=0}^{n-1} \beta_i s^i]/[\sum_{i=-m}^{n} a_i s^i] \qquad (5.4)$$

where

$$\beta_\nu = a_{\nu+1}\gamma_1 + a_{\nu+2}\gamma_1 + \ldots + a_n \gamma_{n-\nu-1} \quad \text{for} \quad \nu = 0, \ldots, n-1 \qquad (5.5)$$

If $n \geq 0$ then $x \in \mathcal{L}$, if $n > 0$ then $x \in \mathcal{C}$.

Proof: Because the inverse of an operator is unique in $\mathcal{O}$ (up to common numerator and denominator factors), Eq. 5.4) defines the only possible solution in $\mathcal{O}$. Because Eq. 5.2) is satisfied as an operator equation by Eq. 5.4) it must be satisfied almost everywhere in R_∞ since the right side is in $\mathcal{L}$ and, hence, also must be the left. The choice of Eq. 5.5) guarantees that the initial conditions of Eq. 5.3) are satisfied.

If $n \geq 0$ we have

$$x = \begin{cases} (s^m f/[\sum_{i=-m}^{n} a_i s^{i+m}]) + ([\sum_{i=0}^{n-1} \beta_i s^{i+m}]/[\sum_{i=-m}^{n} a_i s^{i+m}]); & n > 0 \\ s^m f/[\sum_{i=-m}^{n} a_i s^{i+m}]; & n = 0 \end{cases} \qquad (5.6)$$

where the "denominators" are now polynomial in s. The (common) first term being a convolution of exponentials with f has $n-1$ derivatives absolutely continuous. We can multiply by s^{n-1} to see this, since we still have a convolution with f and Corollary 3.1 then shows how to find the derivatives. As a consequence $x \in \mathcal{C}$ if $n > 0$ (since $s^{(n-1)}$ is) or if $n = 0$ then $x \in \mathcal{L}$.

If $n \leq 0$ we have

$$x = s^m f/[\sum_{i=-m}^{n} a_i s^{i+m}] \qquad (5.7)$$

$$x^{(n-1)} = s^{m+n-1} f/[\sum_{i=-m}^{n} a_i s^{i+m}] \qquad (5.8)$$

Since $n \leq 0$ this latter is again absolutely continuous (as a convolution integral). Q.E.D.

Uniqueness can be alternately proven if $n > 0$ by assuming x_1 and x_2 satisfy Eq. 5.2). Then as $x_1, x_2 \in \mathcal{C}$, $x = x_1 - x_2 \in \mathcal{C}$ satisfies Eq. 5.2 with $f = 0$ and all γ_1 zero. Thus $x = 0$ as it is zero at $t = 0$ and continuous.

The following examples illustrate this method and also clarify the conditions required for a solution.

<u>Example</u> 5.2: The absolute continuity required in condition 1) of Definition 5.1 is needed to give the uniqueness in a functional sense. To see this consider the network of Fig. 5.1.

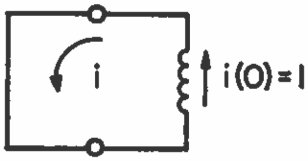

Figure 5.1

Short Circuit Inductor

This has $i' = 0$ with $\{i(0)\} = \{1\}$. Since

$$i'' = si - i(0)$$

we have

$$i = \{1\} = i(0)/s = i(0)\ell$$

Note however that

$$i_1 = \begin{Bmatrix} 1 & 0 \leq t \leq 1 \\ 2 & 1 < t \leq \infty \end{Bmatrix} \in \mathcal{L}$$

has the proper initial value and satisfies $i' = 0$ almost everywhere. Of course i_1 does not operationally satisfy $i_1 = 0$ since

$$si_1 = sH_o + sH_1 = h^o + h^1 = 1 + h^1 = i_1' + i_1(0)$$

(if we consider $sf = f' + f(0)$ to hold here) or

$$i_1' = h^1 \neq 0 \in \mathcal{C}$$

Example 5.3: The network of Fig. 5.2 can not be solved by the Laplace transform

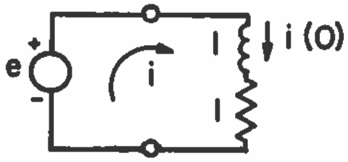

Figure 5.2

Laplace Transform Example

where $e = \{(2t+1)e^{t^2}\}$

$\{i(0)\} = \{2\}$

We have

$$i'+i = \{(2t+1)e^{t^2}\} \text{ with } \{i(0)\} = \{2\}$$

Then

$$i = [2 + \{(2t+1)e^{t^2}\}]/(s+1)$$

$$= \{e^{-t}\}\cdot[2 + \{(2t+1)e^{t^2}\}]$$

or upon performing the indicated convolution

$$i = \{e^{-t} + e^{t^2}\}$$

Note that i can not be written completely in terms of s and that the convolutions required for a solution are sometimes hard to evaluate.

Example 5.4: Consider the network of Fig. 5.3

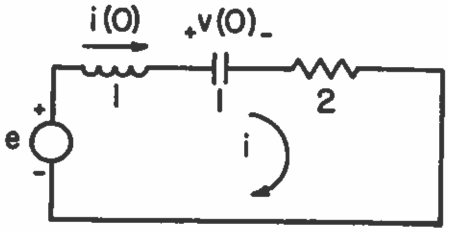

Figure 5.3

Series Circuit for Second Order Equations

where $\{v(0)\} = \{1\},\{i(0)\} = \{0\}$ and e is given by the pulse

$$e = H_o - H_1$$

which is shown in Fig. 5.4.

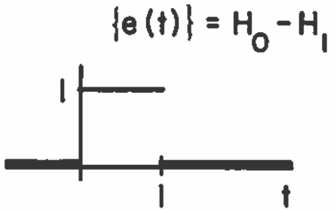

Figure 5.4

Excitation for Fig. 5.3

Then

$$i'+2i+\ell i = e - \{v(0)\}; \quad \{i(0)\} = \{0\} = 0$$
$$\{v(0)\} = \{1\}$$
$$e = H_o - H_1$$

is the describing equation.

This is

$$i = [e - \{v(0)\}]/(s+2+\ell) = [s(H_o-H_1) - s\,\{v(0)\}]/(s+1)^2$$

$$= [h^o - h^1 - v(0)]/(s+1)^2 = [h^o - h^1 - h^o]/(s+1)^2$$

or

$$i = -h^1 \cdot \{te^{-t}\}$$

which is shown in Fig. 5.5.

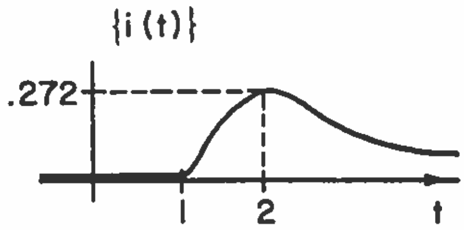

Figure 5.5

Response of Fig. 5.3

Note that even though e is discontinuous i satisfies the equation given for all $t \epsilon R_\infty$. This would not be true, however, if we were to alter $e(t)$ at $t = 1$ (although the new e must be identified with the old in $\mathfrak{L}$). Further note that the initial capacitor voltage must be interpreted as a constant <u>function</u>, not a constant operator.

In systems theory the concept of impulse response is important.

<u>Definition</u> 5.2: The expression

$$x_{imp} = 1/[\sum_{i=-m}^{n} a_i s^i] = s^m/[\sum_{i=-m}^{n} a_i s^{i+m}] \qquad (5.9)$$

<u>is called the impulse response of Eq. 5.2)</u>. Note that Eq. 5.9) is obtained from Eq. 5.4) by assuming $f = 1$, and all $\beta_i = 0$. However, it can not be called a solution of Eq. 5.2) with f replaced by 1 since:

a) Eq. 5.3) can not be satisfied

b) Satisfaction of Eq. 5.2) almost everywhere with $f = 1$ has no meaning.

<u>Example</u> 5.5: Consider the excited capacitor of Fig. 5.6.

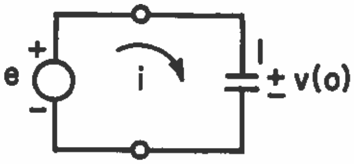

Figure 5.6

Voltage Excited Capacitor

This has

$$\ell i = e - \{v(0)\}$$

and

$$i_{imp} = 1/\ell = s$$

Note that $s \in \mathcal{O}$, $s \notin \mathcal{L}$.

Since $\mathcal{O}$ is a field we can solve multidimensional equations by the use of matrices with entries in $\mathcal{O}$. This is omitted since it is straightforward and familiar from a glance at the corresponding theory using Laplace transforms.

VI. EXTENSIONS AND COMMENTS

In order to consider possible extensions we review very very briefly the theory of distributions [SC 1]. This because the theory of distributions allows an alternate way of defining the impulse and all its derivatives. The use of distributions also allows us to obtain the most general operational calculus.

By a distribution is meant a linear continuous functional defined over the space $\mathfrak{D}$ of testing functions (infinitely continuously differentiable functions which are zero outside a compact set). If we let f be a distribution we can think of it given by $f = < f, >$ where the value of this functional $< f, >$ at the testing function φ is $< f,\varphi >$. Then by definition a distribution is anything with the following two properties

1. $< f, a\varphi_1 + b\varphi_2 > \;=\; a < f,\varphi_1 > + \, b < f,\varphi_2 >$ (linearity)
   for all constants a,b and all $\varphi_1, \varphi_2 \epsilon \mathfrak{D}$
2. $\lim\limits_{n\to\infty} < f,\varphi_n > \;=\; < f, \lim\limits_{n\to\infty} \varphi_n >$ (continuity)
   where all φ_n have their support in a fixed compact set and $\lim\limits_{n\to\infty} \varphi_n$ existing means that φ_n as well as all its derivatives form uniformly convergent sequences.

[Comment: continuity seems somewhat unimportant for physical systems since, although linear discontinuous functionals exist they have never been seen! (to our knowledge)]

Any function in $\mathfrak{L}$ is a distribution since it can be defined by

$$< f,\varphi > \;=\; \int_{-\infty}^{\infty} f(t)\varphi(t)dt \qquad (6.1)$$

But other things, as the impulse δ, are distributions, for instance

$$< \delta,\varphi > \;=\; \varphi(0) \qquad (6.2)$$

The properties of distributions are defined as extensions of those of functions and found by observing the behavior of Eq. 6.1). For example, for any distribution f' is defined by

$$< f',\varphi > \; = \; - < f,\varphi' > \qquad (6.3)$$

(extend Eq. 6.1) by observing its behavior on integrating by parts). This shows any distribution has derivatives of all orders. One can define the convolution by (when it exists)

$$< f*g(t),\varphi(t) > \; = \; < g(t), < f(x),\varphi(xtt) > > \qquad (6.4)$$

and the Laplace transform by

$$\mathfrak{L}[f] \; = \; < f,e^{-st} > \qquad (6.5)$$

when this exists, i.e., when exp[-st] can be considered a testing function. By Eq. 6.2)

$$\mathfrak{L}[\delta] \; = \; < \delta,e^{-st} > \; = \; e^{-s.0} \; = \; 1 \qquad (6.6)$$

and by Eq. 6.5)

$$\mathfrak{L}[f*g] \; = \; \mathfrak{L}[f]\cdot\mathfrak{L}[g] \qquad (6.7)$$

whenever $\mathfrak{L}[f]$ and $\mathfrak{L}[g]$ exist for a common s. In Eq. 6.1) we can consider t to be an n-dimensional variable. Finally, if all components of t are nonnegative and f is zero over the complement of this set of t then we say f has its support in R^n_+ and write $f\epsilon\mathfrak{D}'_+$. When $f\epsilon\mathfrak{D}'_+$ then its convolution with another $g\epsilon\mathfrak{D}'_+$ always exists. In fact $\mathfrak{D}'_+$ forms a commutative ring with no zero divisors (anologous to $\mathfrak{R}_c$) and a unit δ.

The question arises, which is more general, in one dimension, the set of operators $\mathcal{O}$ or the set of distributions $\mathfrak{D}'$. The answer is that neither is contained in either, which is not surprising in that they were defined in completely different ways. For instance the constant function

$\{1\}$ for all time is not an operator, $\{1\} \notin \mathcal{O}$ when $\{1\}$ is defined over R. Further, if $a = \{a(t)\} = \{t^{\mu}\alpha(t)\} \in \mathcal{C}$ and $b = \{b(t)\} = \{\exp(-t^{-2})\} \in \mathcal{C}$ then

$$\theta = a/b$$

is an operator, $\theta \in \mathcal{O}$, which is not a distribution, $\theta \notin \mathfrak{D}'$, for any $\alpha(t)$ with $\alpha(0_+) \neq 0, \mu > 0$ [FE 1, p. 164]. (The equation $\{e^{-1/t^2}\} * \varphi = \underbrace{\{H_o * \ldots * H_o *}_{\text{p times}} \{t^{\mu}\alpha(t)\}\}$ is not solvable for φ for any p).

The theories of distributions and operators are then distinct. One consequently asks if the same results can be accomplished using $\mathcal{L}[f]$ for $f \in \mathfrak{D}'$ as using operators. Again the answer is negative since $\mathcal{L}[e^{t^2} H_o]$ does not exist in the distributional sense.

Because $\mathfrak{D}'_+$, in n-dimensions, is an integral domain one sees that the theory of operators can be extended to n-dimensions by beginning with an n-dimensional $\mathfrak{R}_c$. But one can get a theory which contains the $\mathfrak{D}'_+$ distributions as well as all n-dimensional operators. This results by extending $\mathfrak{D}'_+$ to its field of quotients and has been carried out by Vasilach [VA 2]. Such is the most general theory presently available, but one should be able to work over all of n-dimensional space R^n in a manner analogous to using the bilateral Laplace transform.

As yet no true theory of operators suitable for time-variable equation analysis has been published though much of this is available. However, a similar but somewhat different theory is available for single time-variable equations in terms of 1-dimensional distributions [DO 1], [DO 2].

VII. CRITIQUE

We can now appreciate some of the advantages and disadvantages of this operational calculus of Mikusiński.

Comparing with the Laplace transform we have

A. Advantages:

1) It rigorously defines the impulse and all of its derivatives.
2) It allows solution of more complicated equations.
3) The basic operation of convolution is focused upon with more force than in Laplace transform theory.
4) It appears to extend more naturally to time varying systems since time functions are merely operators.

B. Disadvantages:

1) The interpretation of s as a complex number does not fit naturally into the theory.
2) The concept of frequency response is missing and no reduction to the Fourier integral exists (though these can be obtained by analogy).
3) Finding an operator as a function of s is difficult, as is often the direct evaluation of convolutions.

Compared to the theory of distributions we have

C. Advantages:

1) Basically the concept is simple, no testing functions are needed.
2) It solves the equations of systems theory in a simple way. With distributions we still must use Laplace transforms to get a comparative simplicity.

D. Disadvantages:

1) The main property of the impulse is obscured, that is $\int_{-a}^{a} fh^{o}dt = f(0)$. One does not get the physical picture for the operators that is present with distributions.

2) Most engineers are not very familiar with the algebraic concepts used, as rings and fields, etc.. The distributional notion of a linear functional seems more familiar since it is directly associated with physical measurements.

3) For problems outside the solution of integro-differential equations (e.g., field theory, quantum mechanics, etc.) the operators have not found as many applications.

Almost all disadvantages are overcome by considering imbedding the integral domain of distributions $\mathfrak{D}'_+$ in its field of quotients. This gives an extremely general, complete and beautiful theory.

Epilogue:

Beauty is nothing
but the beginning of terror we can just barely endure,
and we admire it so because it calmly disdains
to distroy us.

Rilke, The First Elegy [RI 1, p. 3]

VIII. ACKNOWLEDGMENTS

The material was developed for two seminars, EE 298 at Berkeley (Spring, 1960 under E. Kuh) and EE 248 at Stanford (Fall, 1964 under T. Kailath). Report preparation has been delayed until now (1967) due to the press of other matters though the concepts are still as pertinent and unavailable as before. The author is indebted to Mary Ellen Terry for her usual care in the preparation of the manuscript.

XI. REFERENCES

[BI 1] Birkhoff, G. and Mac Lane, S., "A Survey of Modern Algebra," Revised Edition, Macmillan, 1965.

[CO 1] Courant, R., "Differential and Integral Calculus," Volume II, English Edition, Interscience, 1952.

[DO 1] Doležal, V., "Über die anwendung von Operatoren in der Theorie der linearen dynamischen Systeme," Aplikace Mathematiky, svazek 6, čislo 1, 1961, pp. 36-67.

[DO 2] Doležal, V., "Über eine Klasse linearer Operatoren," Časopis pro pěstováni mathematiky, svazek 86, čislo 2, 1961, pp. 200-232.

[FE 1] Fenyö, S., "Über den Zusammenhang zwischen den Mikusińskischen Operatoren und Distributionen," Mathematische Nachrichten, vol. 19, nos. 1-6, 1958, pp. 161-164.

[MC 1] McCoy, N., "Rings and Ideals," Mathematical Association of America, 1948.

[MC 2] McShane, E. J., "Integration," Princeton University Press, 1944.

[MI 1] Mikusiński, J. G., "Operational Calculus," Pergamon Press, 1959.

[MI 2] Mikusiński, J. G., "L'anneau algebrique et ses applications dans l'analyse fonctionnelle," Annales Universitatis Mariae Curie--Skłodowska, Lublin, 1st Part, vol. II, no. 1, 1947, pp. 1-47, 2nd Part, vol. III, no. 1, 1949, pp. 1-82. (This contains many extensions of the foregoing theory to other areas of analysis. Part two contains some corrections to part one.)

[RI 1] Rilke, R. M., "Duino Elegies," University of California Press, 1963.

[SC 1] Schwartz, L., "Théorie des distributions," Hermann, vol. I, 1957, vol. II, 1959.

[VA 1] Vasilach, S., "Sur un calcul opérationnel algebrique pour fonctions de deux variables," Revue de Mathématiques Pures et Appliquees, vol. 2, 1957, pp. 181-238 (chapter 1). (This extends Mikusiński's theory to n variables giving examples in 2 variables.)

[VA 2] Vasilach, S., "Calcul opérationnel algebrique des distributions a support dans R_+^n, $n \geq 1$," Revue de Mathématiques Pures et Appliquées, vol. 4, no. 2, 1959, pp. 185-219 (chapter 1), vol. 5, nos. 3-4, 1960, pp. 495-531 (chapter 2), vol. 6, no. 1, 1961, pp. 69-100 (chapter 3). (This extends Mikusiński's theory to the ring of distributions.)