An automatic control course without the Laplace transform

An operational point of view

F. ROTELLA*, I. ZAMBETTAKIS**

* Ecole Nationale d’Ingénieur de Tarbes,
47 avenue d’Azereix, BP 1629, 65016 Tarbes CEDEX, France
rotella@enit.fr

** Institut Universitaire de Technologie de Tarbes,
Université Toulouse III Paul Sabatier,
1 rue Lautréamont, 65016 Tarbes CEDEX, France
izambettakis@iut-tarbes.fr

Abstract— The question is: Is the Laplace transform needed in an Automatic Control course? The answer is: Obviously, not! Based on an operational standpoint the first parts are about guidelines for a primer in automatic control. Beyond an undergraduate course, the two last parts, a little bit more technical, are devoted on the one hand to get the model of a computer controlled system and on the other hand to relate the operational standpoint to usual tables used in some cases.

Résumé— Devant les inconvénients pédagogiques engendrés par l’utilisation de la transformée de Laplace en automatique le développement de méthodes basées sur une présentation purement opérationnelle permet de focaliser les raisonnements sur l’aspect pratique de cette discipline. À partir d’un point de vue développé par Heaviside, nous passons en revue, dans les premières parties, quelques résultats de base en automatisque. Ces résultats sont bien sûr obtenus à l’aide de la notion de transfert ce qui nous permet d’en préciser l’interprétation et les différentes acceptions. Les deux dernières parties, qui ne sont pas nécessaires au premier abord sont consacrées d’une part à la construction du modèle d’un système commandé par calculateur et d’autre part à la liaison que l’on peut établir entre l’approche proposée et l’utilisation des tables. Cet article n’a pas pour objectif de changer l’automatique mais se propose de fournir les moyens d’en changer la présentation actuelle.

Index Terms— Transfer operator, automatic control course, operational calculus, Laplace transform, Carson transform, Heaviside, Mikusiński.

Mots-clés—Opérateur de transfert, cours d’automatique, calcul opérationnel, transformée de Laplace, transformée de Carson, Heaviside, Mikusiński.

I. INTRODUCTION

First level courses in automatic control or basic textbooks of this topic contain first lessons of Laplace transforms. The Laplace transform approach leads to define the transfer function of a system. It is used to get the corresponding response signal of a system with respect to a given input signal. The Laplace transform is also important for the analysis and design of control systems [13]. This tool appears thus a necessary and unavoidable burden for students participating in automatic control courses. However, the transform has some skeletons-in-the-closet [27], [34]. In this article, we discuss an alternative approach to the use of Laplace transform in automatic control courses.

Consider a signal \( x(t) \), defined for a positive time \( t \) and satisfying some appropriate growth conditions. The Laplace transform of \( x(t) \) is

\[
X(s) = \mathcal{L}\{x(t)\} = \int_0^\infty x(t)e^{-st}\,dt, \tag{1}
\]

where \( s \) is a complex variable. This definition can require advanced mathematical machinery [34]. This machinery is very demanding, and usually beyond the skills of most undergraduate students. This generates difficulties that lead desertion of students from automatic control classes. Equation (1) assumes that all considerations, diagrams and developments are embedded in a space of transformed signals. Students ask frequently two questions in regards to the usefulness of equation (1): How can we experimentally exhibit or visualize the transformed signals for example with an oscilloscope? Are there some forbidden signals in automatic control methods? For instance, \( \exp(t^2) \) has no Laplace transform [36].

Some fundamental theorems of Laplace transform have also provided some misunderstandings about the actual meaning of \( s \). For instance, consider the Laplace transform of a derivative function with the initial condition \( x(0) \). Namely, \( \mathcal{L}\{\dot{x}(t)\} = sX(s) - x(0) \). When \( x(0) \) vanishes the complex variable \( s \) is considered as a time derivative operator. However, this can be considered in only the space of transformed signals. Still about this fundamental theorem, the lower limit of integration in equation (1) is often replaced by 0\(^-\), 0\(^+\), or \(-\infty\) [13], [30],
zero initial conditions. In other words, the transfer function is a quotient of the Laplace transform of the output signal. For a single-output system the transfer function is defined as the ratio of the Laplace transform of the output to that of the input, assuming zero initial conditions. In other words, the transfer function is given by:

\[ F(s) = \frac{Y(s)}{U(s)} \]

Although the name transfer function as a mathematical tool is adequate for \( s = j\omega \), where \( j^2 = -1 \) and \( \omega \) is the frequency [23], [3], this ad-hoc definition generates some interesting questions. The Laplace transform of signals cannot be obtained in practice, and sometimes we wonder how to determine the transfer function of a system? For instance this question is avoided in identification procedures [31] where ARMAX models involving recurrence relationships instead of transfer functions are used. In several high quality textbooks for discrete-time systems e.g. [2], the complex variable \( z \) of the Z-transform [26] and the shift-forward operator \( q \) are both used. However, the choice between \( z \) and \( q \) is not argued in every case. These subtle differences, which are mysterious for students, are useless and can be avoided. In view of our personal experience in teaching automatic control, we desire to give an experimental meaning of the transfer of a system, irrespective of previous formal definitions. Rarely, in real occasions students may relate the transfer to the differential operation induced by the system. Indeed the use of the Laplace transform leads to the diagram depicted in Figure 1. This figure describes the relationship between the Laplace transforms of external signals and the transfer of the system. So, the essential meaning is lost. It can be noticed here that we do not use the term transfer function, although we call it transfer as can be seen in the sequel.

As a matter of fact, the use of Laplace transforms is one method among many others [33], [29] to justify the Heaviside operational calculus (Figure 2). Note that the operational calculus is used to solve differential equations (in most cases linear) rather than automatic control problems. From an historical perspective, the Laplace transform, which has been studied extensively by Deakin [8], [9], was introduced first in the form as equation (1) by Bateman in 1910 to solve the differential equation \( \ddot{x}(t) = -\lambda x(t) \) where \( \lambda \) is a nonzero real number.

In the following, we describe guidelines for starting an automatic control course without using the Laplace transform. The presentation is based on the use of a pure operational point of view that provides an opportunity to link methods developed in automatic control to laboratory applications. The essential proofs based on Laplace transform theorems can be read in standard textbooks of automatic control (e.g. [20], [30], [22]). Concerning the notation, we consider signals as elements belonging to the set \( \mathbb{C} \) of integrable real valued functions \( f = \{ f(t) \} \), supposed to be \( m \) times continuously differentiable on \([0, \infty)\) except at isolated points where it is assumed that both left limit and right limit exist. As \( \{ f(t) \} \) denotes the signal \( f \) while \( f(t) \) stands for its value at time \( t \) we write for two signals \( a \) and \( b \) in \( \mathbb{C} \): for all \( t \geq 0 \), \( a(t) = b(t) \), or \( \{ a(t) \} = \{ b(t) \} \), or \( a = b \). However, when no confusion is possible the braces or “for all \( t \geq 0 \)” may be dropped. Nevertheless we do not deal with discrete-time systems except to define the transfer of a computer controlled system.

II. Transfer Operator

We begin by considering a linearized system around an equilibrium point. We suppose this system can be described by the linear differential equation

\[ y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \ldots + a_1y^{(1)}(t) + a_0y(t) = b_0u(t) + b_1u^{(1)}(t) + \ldots + b_mu^{(m)}(t) \]

where \( y(t) \) and \( u(t) \) stand for the differences of output and input signals with their setpoint values respectively and \( n \) and \( m \) are two integers. In (2) the coefficients \( a_i \) and \( b_j \) are constant parameters.

A. Coding

The aim is to provide a tool with which it is easy to manipulate mathematical expressions, which may be used instead of linear differential equations, to separate input and output variables from the system. Following Heaviside [24], [40] or Carson [5], we introduce the derivative operator

\[ p \triangleq \frac{d}{dt} \]

which upon acting on \( x(t) \) in \( \mathbb{C} \) gives the codings

\[ \dot{x}(t) = px(t), \quad \ddot{x}(t) = p^2x(t), \ldots, \quad x^{(n)}(t) = p^nx(t), \ldots \]

In view of these codings and using the distributivity property for every real numbers \( \alpha \) and \( \beta \),

\[ [\alpha p^n + \beta p^m]x(t) = \alpha x^{(n)}(t) + \beta x^{(m)}(t), \]

equation (2) becomes

\[ [p^n + a_{n-1}p^{n-1} + a_{n-2}p^{n-2} + \ldots + a_1p + a_0]y(t) = [b_mp^m + b_{m-1}p^{m-1} + \ldots + b_1p + b_0]u(t). \]
Leibniz (1695) Euler (1730)
Laplace (1812) Servois (1814)

Cauchy (1827) Gregory (1846) Boole (1859)
Heaviside (1884-1895) Kirchhoff (1891)

Wagner (1915) Bromwich (1916) Carson (1917-1922)
Jeffreys-March (1927) Van der Pol (1929) Doestch (1930)
Florin (1934) Schwartz (1945,1947)

B. Operational calculus as polynomial calculus

Operational calculus is understood as algebraic methods for solving differential or recurrence equations, specifically in a linear time-invariant framework. In our point of view solving a differential equation for a given input, is not the aim of automatic control but is a mathematical exercise [34]. In automatic control operational calculus means rules for transfer connections or decompositions through polynomial calculus. Thus the coding (4) is useless when we are not allowed to associate transfer operators. From the operational standpoint the connection rules can be demonstrated through the following steps. The transfers of the connected system provide differential equations. The connections and the elimination of intermediate signals lead to a differential equation between the output and input signals. The encoding of this differential equation with $p$ ensures the results. Although we can use this procedure in every case it is sufficient to exemplify with respect to series or parallel connections for two first-order systems.

Let us consider two linear systems described by $y_1(t) = F_1(p)u_1(t)$ and $y_2(t) = F_2(p)u_2(t)$ where $u_1(t)$ and $u_2(t)$ are the input signals, $F_1(p)$ and $F_2(p)$ the transfers of the systems, and $y_1(t)$ and $y_2(t)$ are the corresponding output signals. In this regard we have

$$F_1(p) = \frac{b_1 p + b_0}{a_1 p + a_0}$$

and

$$F_2(p) = \frac{\beta_1 p + \beta_0}{\alpha_1 p + \alpha_0},$$

where $a_0$, $a_1$, $b_0$, $b_1$, $\alpha_0$, $\alpha_1$, $\beta_0$, and $\beta_1$ are constant parameters. The series connection is defined as $u_2(t) = y_1(t)$,

To separate input and output signals from the system we divide equation (4) by the polynomial $p^n + a_{n-1} p^{n-1} + \cdots + a_0$, which yields to code the input-output relationship as $y(t) = F(p)u(t)$ with

$$F(p) = \frac{b_m p^m + b_{m-1} p^{m-1} + \cdots + b_1 p + b_0}{p^n + a_{n-1} p^{n-1} + a_{n-2} p^{n-2} + \cdots + a_1 p + a_0}. \tag{5}$$

We must insist here that $y(t) = F(p)u(t)$ cannot be considered as the solution of the differential equation (2). Indeed, the initial conditions are not known. $y(t)$ is determined with this writing as with the writing (2). In equation (5), $F(p)$ stands for the transfer operator. In essence it is the transfer, which represents the operation induced by the system to transform the input signal into the output signal. The operational approach provides an opportunity to relate the transfer operator (5) to the differential equation (2). The diagram, depicted in Figure 3 corresponds to an experimental situation.
where we recognize the product \( F_1(p)F_2(p) \). The parallel connection is defined as \( u_2(t) = u_1(t) = u(t) \) and \( y(t) = y_1(t) + y_2(t) \), which yields to

\[
y(t) = \left[ \frac{1}{\alpha_1a_1p^2 + (\alpha_0a_1 + \alpha_1a_0)p + \alpha_0a_0} \right] u(t),
\]

where we recognize the sum \( F_1(p) + F_2(p) \). These results can be extended to any order by induction. It can be seen that the transfer of the series connection of two systems is the product of their transfer and the transfer of the parallel connection of two systems is the sum.

The parallel and series rules give a meaning to the decompositions and the handling of transfer operators with polynomial calculus. These operations on transfer operators are the basis of operational calculus in automatic control. We can apply usual techniques as Mason’s rule associated to the signal-flow graph [35]. This operational calculus can be applied also for multiple-input multiple-output systems with no need of the Laplace transform.

C. The delay operator

A pure time delay of \( T \) between input and output signals induces \( y(t) = u(t - T) \). This particular linear system cannot be associated to a differential equation as (2). A special treatment must be used for delay equations. Following an idea of Euler [14], the Taylor expansion of \( u(t - T) \) yields

\[
u(t - T) = u(t) - \dot{u}(t)T + \frac{\ddot{u}(t)}{2}T^2 - \ldots + u^{(n)}(t)\frac{{(-T)^n}}{n!} + \ldots,
\]

which is encoded to give

\[
y(t) = u(t) - pTu(t) + p^2\frac{T^2}{2}u(t) - \cdots + p^n\frac{{(-T)^n}}{n!}u(t) + \cdots,
\]

\[
= \left( \sum_{n \geq 0} \frac{{(-pT)^n}}{n!} \right) u(t) = (\exp(-pT)) u(t).
\]

We obtained the transfer operator for the time delay \( T \) as \( F(p) = e^{-pT} \).

III. SYSTEM RESPONSES

System analysis is often the study of some particular responses of the system. Of special interest are the step and frequency responses.

A. Step response

Let us consider the Heaviside or step signal \( \{H(t)\} \) defined by \( H(t) = 1 \) for \( t \geq 0 \) and 0 elsewhere. The step response of a system is the solution of the differential equation of the system to a step input signal with zero initial conditions. With operational calculus, we can expand transfer operator as a linear combination of simple transfers

\[
\frac{a^n}{(p + a)^n} \quad \text{or} \quad \frac{1}{p^n},
\]

where \( a \) stands for a nonzero complex number and \( n \) for an integer. The step response of a multiple integrator with the transfer \( \frac{1}{p^n} \) is \( \frac{1}{n!} \). Let us consider the step response \( s_n(t) \) of the Strejč system [41] with the transfer \( \frac{a^n}{(p + a)^n} \). For \( n = 1 \) the associated differential equation to the transfer \( \frac{a}{p + a} \) is \( ay(t) + y(t) = au(t) \) and we obtain by usual methods [51] the corresponding response to a given input \( u(t) \) with the initial condition \( y(0) \)

\[
y(t) = e^{-at} \left( y(0) + a \int_0^t e^{av} u(v) \, dv \right).
\]

For \( u(t) = 1 \) and zero initial conditions the step response becomes

\[
s_1(t) = 1 - e^{-at}.
\]

For \( n = 2 \) we have \( s_2(t) = \frac{a}{p + a} s_1(t) \) from which follows

\[
s_2(t) = ae^{-at} \int_0^t (e^{av} - 1) \, dv = 1 - e^{-at}(1 + at).
\]

For the general case \( n \geq 1 \), let us suppose

\[
s_n(t) = 1 - e^{-at} \left( \sum_{i=0}^{n-1} c_{i,n} t^i \right),
\]

and

\[
s_{n+1}(t) = 1 - e^{-at} \left( \sum_{i=0}^{n} c_{i,n+1} t^i \right),
\]

where the coefficients \( c_{i,n} \) and \( c_{i,n+1} \) are constant parameters. These signals are linked by the differential equation

\[
\dot{s}_{n+1}(t) + as_{n+1}(t) = as_n(t),
\]

which to the the relationships

\[
c_{0,n+1} = 1,
\]

\[
\text{for } i = 1 \text{ to } n : c_{i,n+1} = \frac{a}{k} c_{i-1,n} = \frac{a^i}{i!}.
\]

We deduce the well known result that, for the Strejč model

\[
\frac{a^n}{(p + a)^n},
\]

the step response \( s_n(t) \) is

\[
s_n(t) = 1 - e^{-at} \left( \sum_{i=0}^{n-1} \frac{a^i t^i}{i!} \right).
\]

So the step response of a given system defined by a transfer operator can be calculated using polynomial calculus.
B. Stability

Due to the fact that stability analysis does not use the Laplace transform there is no novelty in this paragraph. However, we may insist again on the link between the transfer operator, the differential equation and the transient behavior.

Let us consider the transfer operator \( \frac{a^n}{(p + a)^n} \) where \( a \) and \( n \) have the same meaning as before. From the previous paragraph we can see that the step response is composed of a constant term and a time-dependent term. The first term is the forced response and the second term is the transient behavior. The transient behavior tends asymptotically to zero if and only if the real part of \( a \) is strictly negative. More generally, consider the \( n \)-th order transfer

\[
F(p) = \frac{b_m p^n + \cdots + b_1 p + b_0}{(p - p_1)^{r_1}(p - p_2)^{r_2} \cdots (p - p_r)^{r_r}},
\]

where \( p_1, p_2, \ldots, p_r \) are the \( r \) complex poles of \( F(p) \) and \( \rho_i \) are the respective multiplicities with \( n = \sum_{i=1}^{r} \rho_i \). For the transient response the poles generate terms associated with the signals \( e^{\rho_1 t}, e^{\rho_2 t}, \ldots, e^{\rho_r t} \) weighted by time polynomials of order \( \rho_i - 1 \) respectively. If all the poles have strictly negative real part, the transient behavior tends asymptotically to zero. Namely, the system is asymptotically stable.

C. Frequency response

For asymptotically stable systems the frequency response is deduced from the steady-state output response corresponding to a given sinusoidal input signal \( u(t) = e^{i\omega t} \) where \( \omega \) is the frequency and \( i^2 = -1 \). Consider a system defined by the transfer operator \( F(p) = \frac{B(p)}{A(p)} \) where \( A(p) \) and \( B(p) \) are two polynomials. From the operational approach we have the encoded input-output differential equation as \( A(p) y(t) = B(p) u(t) \). For the sinusoidal input \( u(t) = e^{i\omega t} \), we obtain

\[
B(p) u(t) = |B(j\omega)| e^{i\omega t} = |B(j\omega)| e^{i\omega t + \text{arg}(B(j\omega))},
\]

where \( |B(j\omega)| \) and \( \text{arg}(B(j\omega)) \) denote the module and the argument of the complex number \( B(j\omega) \) respectively. The output \( y(t) \) is the sum of a particular solution of the differential equation and the general solution of the differential equation without second member. The general solution characterizes the transient response that vanishes in case of asymptotically stable systems. For a particular solution, we look for the steady-state behavior as the form \( y(t) = Y e^{i(\omega t + \varphi)} \) where \( Y \) and \( \varphi \) are constant parameters. Replacing this expression for \( y(t) \) in the differential equation yields

\[
Y = |B(j\omega)| = |F(j\omega)|,
\]

\[
\varphi = \text{arg}(B(j\omega)) - \text{arg}(A(j\omega)) = \text{arg}(F(j\omega)).
\]

The frequency response is defined by the evolution of \( |F(j\omega)|, \text{arg}(F(j\omega)) \) as the frequency \( \omega \) varies from 0 to \(+\infty\). We can notice that \( F(j\omega) \) is the transfer function of the system such as Harris defined it [23]. In our standpoint, this transfer function must not be confused with the transfer operator \( F(p) \). Nevertheless, \( F(j\omega) \) such as a function of the frequency is the only actual transfer function. Graphic representations such as Bode, Black-Nichols, or Nyquist loci may be used to analyze the frequency response [30]. For unstable systems the loci are valid as calculated representations only. But for stable systems, experiments cannot allow to obtain the frequency response.

IV. Analysis

A. Poles and zeros

The names of poles and zeros come from the interpretation of a transfer function \( F(p) \) as a function of a complex variable \( p \). This interpretation is a consequence of the formulation of Laplace transform and it misunderstands the physical meaning of these notions. The consideration of a transfer operator as a coding of a differential equation provides an immediate physical interpretation. Namely, let us consider the transfer operator

\[
F(p) = \frac{p + a}{p + b}, \quad (7)
\]

where \( a \) and \( b \) are constant parameters. In the operational standpoint the transfer (7) corresponds to the input-output differential equation \( y(t) + by(t) = \dot{u}(t) + au(t) \) where \( y(t) \) and \( u(t) \) are the output and input signals. First consider \( u(t) = 0 \) for \( t > 0 \) and a nonzero initial condition \( y(0) \) we obtain \( y(t) = y(0)e^{-bt} \) for \( t > 0 \). Second consider a zero initial condition for the output and the input signal \( u(t) = e^{-at} \) for \( t > 0 \) we obtain \( \dot{u}(t) + au(t) = 0 \) so \( y(t) = 0 \) for \( t > 0 \).

In general poles correspond to signals generated by the system with zero input. Zeros correspond to signals absorbed or blocked by the system. Let us write the transfer operator (5) as

\[
F(p) = k \frac{(p - z_1)^{\nu_1}(p - z_2)^{\nu_2} \cdots (p - z_d)^{\nu_d}}{(p - p_1)^{\rho_1}(p - p_2)^{\rho_2} \cdots (p - p_r)^{\rho_r}},
\]

where \( k = b_m, z_i, i = 1, \ldots, d \) and \( p_i, i = 1, \ldots, r \) are complex numbers, and \( \nu_i, i = 1, \ldots, d \) and \( \rho_i, i = 1, \ldots, r \) are integers. For \( i = 1, \ldots, r, e^{\rho_i t} \) is solution of the coded differential equation

\[
(p - p_1)^{\rho_1}(p - p_2)^{\rho_2} \cdots (p - p_r)^{\rho_r} y(t) = 0,
\]

and for \( i = 1, \ldots, d, e^{z_i t} \) is solution of the coded differential equation

\[
(p - z_1)^{\nu_1}(p - z_2)^{\nu_2} \cdots (p - z_d)^{\nu_d} u(t) = 0.
\]

On the one hand we can use the correspondence between \( e^{\rho_i t} \) and the transfer denominator roots \( p_i \) that characterizes the transient rate in the linear constant parameter framework only. The same remark can be said about the correspondence between \( e^{z_i t} \) and the transfer numerator roots \( z_i \). On the other hand this signal approach for the pole and zeros meaning can be extended to time-varying or nonlinear multivariable systems with an algebraic standpoint [15], [16].

In order to underline and to exemplify the important problem of pole/zero cancellation let us consider the series
connection with the systems:
\[
y(t) = \frac{1}{p-1} u(t),
\]
\[
z(t) = \frac{p-1}{p+1} y(t).
\]
The pole 1 induces, in the transient behavior or in the initial conditions effect for the first system, an \( e^t \) signal. This signal is blocked by the second system, which has 1 as zero. As \( \lim_{t \to \infty} e^t = \infty \), this fact forbids such a connection. Indeed, while the input and output signals are zero, there exists in the system a non observed and non controlled unbounded signal. The conclusion is different if we consider the series connection with the systems:
\[
y(t) = \frac{1}{p+1} u(t),
\]
\[
z(t) = \frac{p+1}{p-1} y(t).
\]

Due to the pole/zero cancellation at -1, \( y(t) \) has an \( e^{-t} \) component that vanishes at \( \infty \). Except during the transient behavior, the pole/zero cancellation is acceptable for asymptotically stable cancelled zeros.

B. DC gain

Let us keep in mind that the transfer operator \( F(p) \) in equation (5) is just a coding of the differential equation (2). In the case of an asymptotically stable system, with the input taking a constant value \( U \), the step response analysis indicates that the output tends to a constant value \( Y \) as \( t \) goes to \( +\infty \) given by the relationship \( a_0 Y = b_0 U \). The ratio \( \frac{Y}{U} \) defines the DC gain of the system \( G_{DC} \). The stability condition implies \( a_0 \neq 0 \), and from (5) we obtain \( G_{DC} = F(0) \).

C. Steady-state error analysis

In all this part systems are supposed to be asymptotically stable, namely the transient behavior vanishes and only the permanent behavior remains. For the reference inputs \( r_i(t) \) defined as, for \( t > 0 \), \( r_i(t) = \frac{t^i}{i!} \), and for \( t < 0 \), \( r_i(t) = 0 \), the corresponding outputs are \( y_i(t) = F(p)r_i(t) \). The input-output error \( \varepsilon_i(t) = r_i(t) - y_i(t) \) is called the system error of order \( i \). Two notions can be pointed out here. First a norm of the instantaneous system error \( \varepsilon_i(t) \) characterizes the system performance. Second the value \( \varepsilon_i(\infty) = \lim_{t \to \infty} \varepsilon_i(t) \) during the permanent behavior characterizes the steady-state error. In basic lecture of automatic control this last notion is usually considered. We detail it according to our formulation, namely without the use of the final value theorem.

A steady-state error of order \( N \) is ensured if \( \varepsilon_i(\infty) = 0 \), for \( i = 0 \) to \( N \), and \( \varepsilon_{N+1}(\infty) \neq 0 \). Consider the transfer operator \( F(p) \) in equation (5) of an asymptotically stable system. The corresponding permanent step response value is given by the DC gain \( \frac{b_0}{a_0} \). It can be seen that \( \varepsilon_0(\infty) = 0 \) if and only if \( b_0 = a_0 \). We can conclude that a steady-state error of zero order is fulfilled whether the DC gain is equal to 1. In other words since the input-error transfer is \( 1 - F(p) \), we obtain a steady-state error of zero order when the input-error DC gain is zero. This is a fundamental remark for the following.

If we notice that \( r_1(t) \) is the integral of \( r_0(t) \), namely
\[
r_1(t) = \frac{1}{p} r_0(t),
\]
we have
\[
\varepsilon_1(t) = r_1(t) - y_1(t),
\]
\[
= \frac{1}{p} r_0(t) - F(p) \frac{1}{p} r_0(t),
\]
\[
= \frac{1 - F(p)}{p} r_0(t).
\]

Clearly, from the previous result for \( \varepsilon_0(\infty), \varepsilon_1(\infty) \) vanishes if and only if the DC gain of the transfer operator \( \frac{1 - F(p)}{p} \) is equal to zero. Since
\[
\frac{1 - F(p)}{p} = \frac{(a_0 - b_0) + (a_1 - b_1)p + (a_2 - b_2)p^2 + \cdots}{p(a_0 + a_1p + a_2p^2 + \cdots + a_np^n)}
\]
we obtain \( \varepsilon_1(\infty) = 0 \) if and only if \( a_0 = b_0 \) and \( a_1 = b_1 \). It can be seen that
- when \( a_0 \neq b_0 \), we have \( \varepsilon_0(\infty) \neq 0 \) and with \( \varepsilon_1(\infty) = \lim_{p \to -a_0} \frac{a_0 - b_0}{p(a_0 + a_1p + a_2p^2 + \cdots + a_np^n)} = \pm \infty \);
- when \( a_0 = b_0 \), we obtain \( \varepsilon_0(\infty) = 0 \) and with \( \varepsilon_1(\infty) = \frac{a_1 - b_1}{a_0} \). Moreover \( \varepsilon_1(\infty) = 0 \) when \( a_1 = b_1 \).

In the same way we can show by recurrence that the system can have a steady-state error of order \( N \) if and only if its transfer function \( F(p) \) is such that, for \( i = 0 \) to \( N \), \( a_i = b_i \). In this case the steady-state error of order \( N + 1 \) is
\[
\varepsilon_{N+1}(\infty) = \frac{a_{N+1} - b_{N+1}}{a_0},
\]
and the next ones have an infinite module. The degree of the steady-state error can be obtained by just a visual inspection of the transfer operator of the system.

V. Computer controlled systems

The last question we wish to deal with concerns the construction of the model of a linear system controlled by a numerical algorithm (e.g. [2], [19]). The problem is to find the discrete-time model corresponding to the structure presented in Figure 4 where DAC denotes a digital-analog converter and it is usually modelled as the series connection of an ideal sampler and a zero-order hold. ADC denotes an analog-digital converter usually modelled by an ideal sampler and the discrete output signal of the ADC device is \( y_k = y(kT_s) \) for \( k \) in \( \mathbb{N} \). Both are supposed to be synchronized with the sampling period \( T_s \). In this section we use the notations \( \{x_k\} \) for the discrete-time real valued signal defined for \( k \) in \( \mathbb{N} \) and \( q^{-1} \) for the delay operator \( q^{-1}\{x_k\} = \{x_{k-1}\} \). All the considered discrete-time signals are supposed to be zero for negative values of \( k \). For the discrete input signal \( \{u_k\} \) the output of the DAC device is
\[
u(t) = \sum_{k \geq 0} u_k (H(t - kT_s) - H(t - (k + 1)T_s)),
\]
where $H(t - kT_s)$ stands for the delayed step signal.

The discrete-time transfer of the system (Fig. 4) is obtained through the formula

$$F(q^{-1}) = (1 - q^{-1})Z \left\{ \left[ L^{-1} \left\{ \frac{F(p)}{p} \right\} \right] \right\}, \quad (9)$$

where $L^{-1}$ stands for the inverse Laplace transform and $Z \{ . \}$ stands for the Z-transform [26]. The expression in the square bracket denotes the sampling of the signal with a period $T_s$. Although the Z-transform doesn’t suffer the same drawbacks as the Laplace transform, for instance discrete impulse is really a discrete signal, the Laplace transform appears once more here. To carry on we obtain the discrete-time transfer operator $F(q^{-1})$ in the operational framework.

Denoting by $S(t)$ the step response of $F(p)$, the response of $F(p)$ to $u_k(H(t - kT_s) - H(t - (k + 1)T_s))$ is $u_k(S(t - kT_s) - S(t - (k + 1)T_s))$. The sampling of this response at the time period $T_s$ leads to a value for $t = lT_s$, $l$ in $\mathbb{N}$, given by $u_k(S_{l - k} - S_{l - k - 1})$ where $S_k$ stands for $S(kT_s)$. Denoting $l$ the independent integer variable the corresponding discrete signal is $\{ u_k(S_{l - k} - S_{l - k - 1}) \} = (1 - q^{-1}) \{ u_kS_{l - k} \}$ for a given $k$. Because of linearity we deduce from (8) that the sampled response corresponding to the input signal $\{ u_k \}$ is

$$\{ y_l \} = (1 - q^{-1}) \left\{ \sum_{k=0}^{l} u_kS_{l-k} \right\}.$$

Denoting $\{ \sum_{k=0}^{l} u_kS_{l-k} \} = \{ v_l \}$, $\{ u_k \}$ and $\{ v_k \}$ are linked by a discrete convolution operation. Consequently [30], they are linked by a difference equation such as

$$v_k + d_1 v_{k-1} + \cdots + d_{n'} v_{k-n'} = n_0 u_k + n_1 u_{k-1} + \cdots + n_{m'} u_{k-m'},$$

where the $d_i$ and the $n_i$ are real numbers and $n'$ and $m'$ are integers. In the same idea that for continuous systems, this equation can be coded by means of the delay operator $q^{-1}$

$$(1 + d_1 q^{-1} + \cdots + d_{n'} q^{-n'}) \{ v_k \} = (n_0 q^{-1} + \cdots + n_{m'} q^{-m'}) \{ u_k \},$$

which leads to the discrete transfer operator

$$G(q^{-1}) = \frac{n_0 + n_1 q^{-1} + \cdots + n_{m'} q^{-m'}}{1 + d_1 q^{-1} + \cdots + d_{n'} q^{-n'}}.$$

Let us denote the numerator and the denominator of $G(q^{-1})$ by $N(q^{-1})$ and $D(q^{-1})$ respectively. The division of $N(q^{-1})$ by $D(q^{-1})$ leads to

$$G(q^{-1}) = \sum_{l \geq 0} g_l q^{-l},$$

where $g_l$, $l \geq 0$, are real numbers. With $\{ v_k \} = G(q^{-1}) \{ u_k \}$ we obtain

$$\{ v_l \} = \left\{ \sum_{k=0}^{l} g_k u_{l-k} \right\} = \left\{ \sum_{k \geq 0} u_k S_{l-k} \right\}.$$

Since for $k < 0$, $u_k = S_k = 0$, we have for $k \geq 0$, $g_k = S_k$. So, $G(q^{-1})$ is directly obtained from the sampling of the step response of the continuous-time transfer operator. Finally, we conclude that the discrete-time transfer operator is defined by

$$\{ y_k \} = F(q^{-1}) \{ u_k \} = (1 - q^{-1}) G(q^{-1}) \{ u_k \}.$$

As an example let us consider the first order model

$$F(p) = \frac{K}{1 + \tau p}, \quad (10)$$

where $K$ and $\tau$ are real numbers. The sampling, with period $T_s$, of the step response of $F(p)$ gives

$$g_k = K(1 - D^k),$$

where $D = e^{-\tau T_s}$. We deduce

$$G(q^{-1}) = K \sum_{k=0}^{l} (1 - D^k) q^{-k},$$

$$(1 - q^{-1})(1 - D q^{-1}),$$

so the discrete-time transfer of (10) is

$$F(q^{-1}) = \frac{(1 - D) q^{-1}}{1 - D q^{-1}}.$$

VI. WHAT’S ABOUT TABLES?

The main reason of using the Laplace transform is the tables we have at our disposal. First they contain information to determine the response of a system with respect to a given input signal. Second to obtain the discrete transfer operator of a computer controlled system with the formula (9). Although transforms are not used in our presentation, we show that these tables can be used without any change. To do that, let us introduce the notion of generator of a continuous signal, which consists in writing the time expression of this signal by means of the operator $p$.

The previous parts show that the transfer operator allows to link input $u(t)$ and output $y(t)$ signals of a linear system by a differential equation coded as $y(t) = F(p)u(t)$. Until now we obtained the step response by solving this differential equation when initial conditions are all zero. In case of no input and non zero initial conditions such a transfer operator produces an output signal solution of the associated homogeneous differential equation. The coding of this differential equation with the $p$ operator defines then the generator of this signal. Two ways can be considered to take into account initial conditions in this coding. Namely, on the one hand the Mikusiński operational calculus and on the other hand the integral form of a differential equation.
A. The Mikusiński operational calculus

All the previous developments can be rigorously proved by means of operational calculus of Mikusiński [37], which is based on convolution algebra of operators. Let us briefly describe this operational calculus whereas keeping in mind that the considerations below are not needed in a first level course. Convolution product is a fundamental tool in dynamic systems field [45], [46], specifically in case of linear systems [21], [10]. This tool is defined by

\[ (f, g) \mapsto gf = \int_0^t f(\tau)g(t - \tau)d\tau, \]

while the Heaviside function \( H = \{1\} \) is of great importance due to the fact that we have for every \( f \) in the set of integrable function \( C \)

\[ Hf = \left\{ \int_0^t f(x)dx \right\}. \]

Consequently, \( H \) appears as the integration operator. The successive powers of \( H \) with respect to the convolution product are

for all \( n \in \mathbb{N}, n \geq 1, H^n = \left\{ \frac{t^{n-1}}{(n-1)!} \right\} \).

To distinguish a constant signal \( \{\alpha\} \) with the operator defined by the constant gain \( \alpha \) we denote it by \([\alpha]\). For all \( f \) in \( C \)

\[ \{\alpha\} f = \left\{ \alpha \int_0^t f(\tau)d\tau \right\} \quad \text{and} \quad [\alpha] f = \{\alpha f(t)\} . \]

[1] is the unit element for the convolution and we can give a meaning to \( H^0 \) as \( H^0 = [1] \). We define the derivative operator as the solution of the convolutional equation \( pH = [1] \), and then we write \( p = H^{-1} \). With the understanding \( p^0 = H^{-0} = [1] \), we have \( p^n = H^{-n} \) for \( n \in \mathbb{N} \).

Mikusiński [37] has proved the two results below, which are essential for our purpose.

**Theorem 1** For every continuous function \( f \) in \( C \), \( \{f^{(1)}(t)\} = pf - [f(0)] \). More generally, for every integer \( k \)

\[ \{f^{(k)}(t)\} = p^k f - \sum_{i=0}^{k-1} \left[ f^{(i)}(0) \right] p^{k-i-1} \] (11)

**Theorem 2** For every \( f \) in \( C \) such that \( \int_0^\infty e^{-tp} f(t)dt \) exists

\[ \int_0^\infty e^{-tp} f(t)dt. \]

The first theorem allows to write the generator of a signal \( \{f(t)\} \) when is known the differential equation whose this signal is solution. Indeed, let us suppose that this differential equation can be written

\[ \sum_{i=0}^{n} \alpha_i f^{(i)}(t) = 0, \] (12)

with initial conditions \( f(0) = f_0, f(0) = f_1, \ldots, f^{(n-1)}(0) = f_{n-1} \), where \( n \) is an integer and the \( \alpha_i \) are real numbers. With (11) the coding of (12) leads to

\[ \sum_{i=0}^{n} \alpha_i p^i f(t) - P_{IC}(p, f_0, \ldots, f_{n-1}) = 0, \]

where \( P_{IC}(p, f_0, \ldots, f_{n-1}) \) is a polynomial in \( p \) that depends on the initial conditions and the coefficients \( \alpha_i \). We obtain then the generator of \( \{f(t)\} \)

\[ \{f(t)\} = \frac{P_{IC}(p, f_0, \ldots, f_{n-1})}{\left[ \sum_{i=0}^{n} \alpha_i p^i \right]}. \] (13)

The second theorem indicates that when the one-sided Laplace transform of a signal exists, its expression is identical to the generator of the signal, the complex variable \( s \) of Laplace transform being changed into the derivative operator \( p \) (to avoid any confusion). A major consequence is that the tables [12], [48], can be used. Since \( H = p^{-1} \) we remark that the generator can be written indifferently with the operators \( H \) or \( p \).

For example when we look for a generator, say for \( \sin(\omega t) \), we observe that \( \sin(\omega t) \) is the solution of the differential equation

\[ \ddot{x}(t) + \omega^2 x(t) = 0, \quad x(0) = 0, \quad \dot{x}(0) = 1. \] (14)

Using (11) the coded form is obtained as

\[ p^2 x(t) - 1 + \omega^2 x(t) = 0, \]

which leads to the generator of \( \sin(\omega t) \)

\[ \{\sin(\omega t)\} = \frac{1}{p^2 + \omega^2}, \]

where the symbol “M” denotes “in the Mikusiński sense”. Indeed, we see in the next section an alternative to the Mikusiński approach. The Mikusiński operational calculus has been extended recently by the convolutional calculus [11].

B. The integral form

A second point of view can be used to introduce initial conditions in differential equations. In particular, let us consider the result below [44]

\[ \dot{x}(t) = f(t), \quad x(0) = x_0 \]

if and only if

\[ x(t) = x_0 + \int_0^t f(\tau)d\tau. \] (15)

Denoting the integral operator by \( H \), the derivative operator by \( p \), and using the fact that for zero initial conditions, \( pH = Hp = I \) where \( I \) stands for the identity operator, we can code the differential equation (15) by \( x(t) - x_0 = Hf(t) \) or \( px(t) - px(0) = f(t) = \dot{x}(t) \). In general, from the differential equation

\[ x^{(n)}(t) = f(t), \]

with the initial conditions \( x(0) = x_0, \dot{x}(0) = x_1, \ldots, x^{(n-1)}(0) = x_{n-1} \), we obtain

\[ x(t) = \sum_{k=0}^{n-1} x_k \frac{t^k}{k!} + \int_0^t \int_0^{t} f(\tau)d\tau dt. \] (16)
In particular for \( \{ f(t) \} = 0 \) and the initial conditions \( x_0 = x_1 = \cdots = x_{k-1} = x_{k+1} = \cdots = x_{n-1} = 0 \) and \( x_k = 1 \) we obtain \( x(t) = \left\{ \frac{t^k}{k!} \right\} \). As the corresponding differential equation can be coded \( p^k x(t) = 1 \) we get

for every integer \( k \), \( \left\{ \frac{t^k}{k!} \right\} = \frac{1}{p^k} \).

Moreover for zero initial conditions and for all integer \( k \) we have \( p^k H^k = H_k p^k = I \). These remarks induce for (16) the coding

\[
x^{(n)}(t) = p^n x(t) - \sum_{k=0}^{n-1} x_k p^{n-k}.
\]

The generator of a signal \( \{ f(t) \} \) can be obtained in this framework by coding with (17) the differential equation (12) whose this signal is solution. We obtain the coding

\[
\left[ \sum_{i=0}^{n} \alpha_i p^i \right] f(t) - P_{IC}^*(p, f_0, \ldots, f_{n-1}) = 0,
\]

where \( P_{IC}^*(p, f_0, \ldots, f_{n-1}) \) is a polynomial in \( p \) that depends only on the initial conditions and the coefficients \( \alpha_i \). We get the generator of \( \{ f(t) \} \) as

\[
\{ f(t) \} = \frac{P_{IC}^*(p, f_0, \ldots, f_{n-1})}{\sum_{i=0}^{n} \alpha_i p^i}.
\]

However, if for a given function \( \{ f(t) \} \) we compare the generators (13) and (18) it can be seen that \( P_{IC}^*(p, f_0, \ldots, f_{n-1}) = p P_{IC}(p, f_0, \ldots, f_{n-1}) \). We have the following relationship between the generators: the second one is the first one multiplied by \( p \). The generator (18) is the generator of \( f(t) \) in the Carson sense. The Carson transform was introduced in 1926 [6] and it differs from the Laplace transform by a factor \( p \). The Carson tables can be used in this framework.

For example when we evaluate this generator for \( \sin(\omega t) \), we obtain from the differential equation (14) the coding

\[
p^2 x(t) - p + \omega^2 x(t) = 0,
\]

which leads to the generator of \( \sin(\omega t) \)

\[
\sin(\omega t) = \frac{p}{c} \frac{\sqrt{p^2 + \omega^2}},
\]

where the symbol “\( C \)” denotes “in the Carson sense”. It is obvious that the integral form approach appears more intuitive than the Mikusiński approach but the results are quite similar.

**C. Consequences**

The generator of a signal allows to perform both of the two points below. First, for a system defined by the transfer operator \( F(p) \) we can calculate the response \( y(t) \) to an input \( u(t) \) by using Carson or Laplace transform tables. Indeed when \( U(p) \) is a generator of \( u(t) \) we obtain the generator of the output

\[
y(t) = F(p) E(p).
\]

We must remark here that the generators can be obtained in any sense as defined above (Mikusiński or Carson). However, we must keep the coherence in using tables. For instance, when we want to know the beginning of the response we can develop \( y(t) \) in power of \( p^{-1} \). This procedure was used by Heaviside [24]. However, different functions may be associated to \( p^{-k} \) according to the adopted generator sense. Second, we can obtain the discrete-time model of a computer controlled system. Since we see it above, this model uses the step response \( S(t) \) of a system defined by the transfer operator \( F(p) \). Incidentally, we can remark that the step generator in the Mikusiński sense for \( S(t) \) is \( \frac{F(p)}{p} \). In this framework we obtain the formula (9) for the discrete-time model. Whatever the used generator sense the tables can be used also.

Moreover, in order to see the importance of the generator for operational calculus, let us consider the following example where two signals \( y_1 \) and \( y_2 \) are defined by the differential equations:

\[
(p - 1) y_1(t) = u(t),
\]

\[
(p - 1) y_2(t) = u(t),
\]

and the initial conditions \( y_1^0 \) and \( y_2^0 \) respectively. If we consider the operational calculus, the parallel connection \( y(t) = y_1(t) - y_2(t) \) yields to:

\[
y(t) = \frac{1}{p - 1} u(t) - \frac{1}{p - 1} u(t) = 0.
\]

This conclusion is obviously wrong. Indeed our setting and the suggested proof of the operational calculus indicate that we consider formal differential equations, namely without initial conditions. When we write:

\[
y(t) = \frac{1}{p - 1} u(t),
\]

it is just a coding of the differential equations (19) and (20) only. The initial conditions can be taken into account by mean of the generator notion. We can write (19) and (20) as, respectively:

\[
y_1(t) = \frac{1}{M} u(t) + \frac{y_1^0}{p - 1},
\]

\[
y_2(t) = \frac{1}{M} u(t) + \frac{y_2^0}{p - 1},
\]

in the Mikusiński’s generator sense. It yields for \( y(t) = y_1(t) - y_2(t) \) the generator:

\[
y(t) = \frac{y_1^0 - y_2^0}{p - 1}.
\]

This point indicates that \( y(t) \) is solution of the differential equation:

\[
(p - 1) y(t) = 0, \quad y(0) = y_1^0 - y_2^0,
\]

or, equivalently, \( y(t) = (y_1^0 - y_2^0) e^t \).

This standpoint can also be explained in a more algebraic framework as the Fliess’module-theoretic approach [17]. Nevertheless, let us remind a sentence of a recent paper [18] where is used this point of view for the design of a new identification procedure: “Let us add we tried to write the examples in such a way that they might be grasped without the
necessity of reading the sections on the algebraic background. Our standpoint on parametric identification should therefore be accessible to most engineers.”

VII. Conclusion

We show in this short survey that from the use of the differential operator we obtain all the usual results derived by means of the Laplace formulations. To teach a basic lecture in automatic control using the operational method offers some advantages. The integral or derivative operators allow to link every notion to its physical meaning. We keep in mind that a transfer operator is always related to only a differential equation or a difference equation. Mathematical background is minimized, however, when a rigorous justification is needed the Mikusinski operational calculus may be used. This operational calculus is based on the convolution operator, which is a natural tool for linear equations.

Moreover, we meet here through a pedagogical step the operational standpoints adopted directly in some advanced textbooks to modelize the input-output relationship induced by a linear system. For instance, the discrete-time autoregressive moving-average model

\[ A(q^{-1})y_k = B(q^{-1})u_k + C(q^{-1})\epsilon_k \]

where \( A(q^{-1}), B(q^{-1}), \) and \( C(q^{-1}) \) are polynomials in the delay operator and \( \{\epsilon_k\} \) a noise signal is used in [7], [1] for identification purposes to describe the difference equation between the input \( u_k \) and the output signals \( y_k \) of a given system. For multivariable continuous-time linear systems [42], [50], [49] introduce the model

\[
P(p)\xi(t) = Q(p)u(t),
\]

\[
y(t) = R(p)\xi(t) + W(p)u(t),
\]

where \( P(p), Q(p), R(p), \) and \( W(p) \) are matric polynomials in the differential operator and \( \xi(t) \) is a vector-valued function of time called the partial state. More recently, [4] define the generator of a multivariable system as the polynomial matrix \( M(p) \) in the derivative operator, which allows to write the relationship between input and output signals as

\[
M(p)\begin{bmatrix} y(t) \\ u(t) \end{bmatrix} = 0.
\]

The generator in the sense defined in [4] must not be confused with signal generators. The interested reader can see the quoted literature.

However, the operational pedagogic point of view has one drawback. When one of our students meets an automatic control engineer or reads an automatic control textbook, the student must be warned that he or she might find some not up-to-date operational thoughts using the Laplace transform. But, this gap will be filled in the future.

REFERENCES