Sahand Communications in Mathematical Analysis (SCMA) Vol. 11, No. 1 (2018), 115-132
http://scma.maragheh.ac.ir
DOI: 10.22130/scma.2018.63445.240

The Integrating Factor Method in Banach Spaces

Josefina Alvarez[1], Carolina Espinoza-Villalva[2], and Martha Guzmán-Partida[3*]

Abstract. The so called integrating factor method, used to find solutions of ordinary differential equations of a certain type, is well known. In this article, we extend it to equations with values in a Banach space. Besides being of interest in itself, this extension will give us the opportunity to touch on a few topics that are not usually found in the relevant literature. Our presentation includes various illustrations of our results.

1. Introduction

Most of the undergraduate textbooks on ordinary differential equations, include a section dedicated to the so called integrating factor method (see, for instance, [3, p. 126]), that produces a formula for the general solution of ordinary differential equations of the form

$$\frac{dy}{dt} = u(t)\, y + v(t)\,. \tag{1.1}$$

Here, t is in some interval $[a, b]$ of the real line where it is assumed that the real functions $u(t)$ and $v(t)$ are continuous.

The fundamental idea of the method is to find a function, $\mu(t)$, with the property

$$\mu(t)\left(\frac{dy}{dt} - u(t)\, y\right) = \frac{d}{dt}(\mu y)\,. \tag{1.2}$$

Let us observe that knowing such a function would allow us to write

$$\frac{d}{dt}(\mu y) = \mu(t)\, v(t)\,.$$

2010 *Mathematics Subject Classification.* 46B25, 28B05.
Key words and phrases. Banach spaces, Cauchy-Riemann integral, Exponential function.
Received: 05 May 2017, Accepted: 10 June 2018.
* Corresponding author.

That is to say, the general solution of (1.1) will be the family of antiderivatives of the function $\mu(t)\,v(t)$, each of them divided by $\mu(t)$ and defined where this division makes sense.

It is clear then that knowing a function $\mu(t)$ with the property (1.2) allows us to solve, or integrate, (1.1). This is why $\mu(t)$ is called an integrating factor.

We are left now to figure out how to find such a function. We begin by writing (1.2) as

$$\mu\frac{dy}{dt} - \mu u(t)\,y = \frac{d\mu}{dt}y + \mu\frac{dy}{dt},$$

or

$$\frac{d\mu}{dt}y = -\mu u(t)\,y.$$

So it would suffice to find a solution of

$$\frac{d\mu}{dt} = -\mu u(t)\,,$$

which can be written as

$$\mu(t) = e^{-\int u(t)dt},$$

where $\int u(t)\,dt$ indicates one of the antiderivatives of the function $u(t)$, defined on $[a,b]$. Finally,

$$y(t) = e^{\int u(t)dt}\left(\int e^{-\int u(t)dt}v(t)\,dt + C\right), \tag{1.3}$$

for any real constant C, will be the general solution of (1.1), defined on $[a,b]$. It is possible that some of the antiderivatives involved in (1.3) might not have an explicit form.

Having the presented the method in its usual context, we now wish to extend it to an equation of the same form, under the following assumptions:

(i) The function $u(t)$ is continuous from $[a,b]$ to $L(X)$, the linear and continuous operators from a real Banach space X into itself, with its usual structure of real Banach spaces (see, for instance, [10], Chapter 5). Let us recall that the continuity of a function with values in a Banach space, is defined in the same way as in the case of a real function, only replacing the absolute value, with the norm in the indicated Banach space.

(ii) The function $v(t)$ is continuous from $[a,b]$ to the real Banach space X mentioned above.

We write the vector equation as

$$\frac{dy}{dt} = (u(t))\,(y(t)) + v(t)\,, \tag{1.4}$$

where the unknown function $y(t)$ must be a continuous function from some real intervals to X, with a continuous derivative. As in the case of a real valued function, the derivative is defined as the limit, in the pertinent Banach space, of the ratio of increments. Let us remark that for each t, $(u(t))(y(t))$ is the action of the operator $u(t)$ on the vector $y(t)$.

Formally, we could repeat all the steps described in the scalar case, arriving at the same formula (1.3), which we now write as

$$y(t) = \left(e^{\int u(t)dt}\right)\left(\int \left(e^{-\int u(t)dt}\right)(v(t))\,dt + C\right), \tag{1.5}$$

where the constant should be any vector in X. We would want the integrating factor $e^{-\int u(t)dt}$ to be a continuous function, with one continuous derivative, defined on the interval $[a,b]$ with values in $L(X)$. So, our first order of business is to justify each of the operations involved in (1.5).

Before all else, we need antiderivatives of the functions $t \to u(t)$ and $t \to -u(t)$. If the concepts of Calculus still apply to the vector case, these antiderivatives should be integrals. So, we need to define, in some sense, the integral of a function with values in a real Banach space. Since all the functions involved are continuous, it will be unnecessary to invoke advanced integration methods, such as the Bochner integral (see, for instance, [6, pp. 78-89]), which is based on the Lebesgue integral. As, we will see in the next section, an integral à la Cauchy-Riemann will serve us well.

2. A Cauchy-Riemann Integral With Values in a Real Banach Space

Let us begin by observing that the function $v(t)$ in (1.4), being continuous on the compact set $[a,b]$, is uniformly continuous and bounded. The uniform continuity is proved in the same manner as in the case of a function with real values (see, for instance, [5, p. 95, Theorem 3.8]), replacing the absolute value with the norm in the Banach space X. As for the boundedness, if we follow the proof, for instance in [5, p. 96, Theorem 3.9], a small modification will make it work in the vector case. In fact, we already know that $v(t)$ is uniformly continuous. In particular, there is $\delta > 0$ such that $|t-s| < \delta$ implies $\|v(t) - v(s)\|_X < 1$. We now subdivide the interval $[a,b]$ in a finite number of non overlapping intervals, that is to say, having no more than one point in common,

$$[s_0, s_1], [s_1, s_2], \ldots, [s_{n-1}, s_n], \tag{2.1}$$

with $s_0 = a$, $s_n = b$ and length $< \delta$. Then, if $t \in [a,b]$, we must have $t \in [s_{j-1}, s_j]$ for some j, so $\|v(t) - v(s_j)\|_X < 1$. That is to say, $\|v(t)\|_X < 1 + \|v(s_j)\|_X \leq 1 + \max_j \|v(s_j)\|_X$, which shows that the function $v(t)$ is bounded. Of course, these properties hold for any continuous function from $[a, b]$ into any Banach space.

We are ready now to build the Cauchy-Riemann integral of a continuous function defined on a real interval $[a, b]$ with values on an arbitrary real Banach space, say X, which, when necessary, will be the space $L(X)$ of operators. The same method presented in any Calculus book will give us such an integral. However, as is very well argued, for instance in [1, pp. 330], the nature of the limiting process involved is not quite straightforward. So, we will take this opportunity to sketch the construction of the integral in a rigorous manner, using a few topological tools. Of course, the construction will work *mutatis mutandis* when the function takes real values.

We begin with the following definition:

Definition 2.1 ([8], p. 79)**.** A subdivision S of the interval $[a, b]$ is a finite family of closed non overlapping subintervals covering the interval. With $|I|$ we denote the length of an interval I in the subdivision. The mesh of the subdivision, denoted by $\|S\|$, is $\max_{I\in S} |I|$. For instance, the subintervals in (2.1) are a subdivision with mesh $< \delta$.

We indicate by $\mathfrak{S}$ the family of all the possible subdivisions of $[a, b]$. Finally, to each $S \in \mathfrak{S}$ we associate an arbitrary function $c : S \to [a, b]$, such that $c(I) \in I$, which we will call it the tag of S.

For a fixed function $f : [a, b] \to X$, we can now write down the Riemann sum relative to a subdivision S with tag c as

$$R(S; c) = \sum_{I\in S} |I| f(c(I)) . \tag{2.2}$$

We will show that, when f is continuous, the sums $R(S; c)$ will converge, in some sense, to a unique vector in X, the sense of this convergence being the so called Moore-Smith convergence, which we now explain, in a very simplified version adapted from [8]. For a more in depth discussion, see [2] and the references therein.

Let us recall that a sequence is a function defined on the set of natural numbers. Since $\{R(S; c)\}_S$ is a function defined on $\mathfrak{S}$, with values in X, to talk about convergence we need to have in $\mathfrak{S}$ a version of "for all $n \geq N$". This is done by defining in $\mathfrak{S}$ a relation, denoted $\geq$, as follows:

Definition 2.2. Given two subdivisions $S, S' \in \mathfrak{S}$, we say that $S \geq S'$ if S is a refinement of S'. That is to say, if each subinterval of S is contained in a subinterval of S'.

This relation is reflexive ($S \geq S$), transitive ($S \geq S'$ and $S' \geq S''$ imply $S \geq S''$) and, moreover, it satisfies the following property:

Given two subdivisions $S', S'' \in \mathfrak{S}$, there is a subdivision $S \in \mathfrak{S}$ such that $S \geq S'$ and $S \geq S''$.

A relation $\geq$ with these three properties is said to direct the set $\mathfrak{S}$. The pair $(\mathfrak{S}, \geq)$ is then called a directed set, allowing us to think of the sums $\{R(S;c)\}_S$ as a generalized sequence or net.

Definition 2.3. The net $\{R(S;c)\}_S$ is a Cauchy net in X if, for each $\varepsilon > 0$, there is $S_\varepsilon \in \mathfrak{S}$ so that

$$\left\| R(S;c) - R\left(S';c'\right)\right\|_X < \varepsilon,$$

for all $S, S' \geq S_\varepsilon$ and for all the tags c and c'.

The net $\{R(S;c)\}_S$ converges to $A \in X$ if, for each $\varepsilon > 0$, there is $S_\varepsilon \in \mathfrak{S}$ so that

$$\| R(S;c) - A\|_X < \varepsilon,$$

for all $S \geq S_\varepsilon$ and for all the tags c.

As a direct consequence of this definition, when the vector A exists, it is uniquely determined.

Theorem 2.4 ([8], p. 193, Theorem 24)**.** *In a complete metric space, that is to say a metric space where every Cauchy sequence converges, it is also true that every Cauchy net converges.*

So, to prove that the net $\{R(S;c)\}_S$ given by (2.2) converges in X, when the function f is continuous, it will suffice to show that it is Cauchy. If this is the case, the limit of the net will be denoted by $\int_a^b f(t)\,dt$ and we will say that the integral has been constructed by refinement.

Proposition 2.5. *If $f : [a,b] \to X$ is continuous, the net $\{R(S;c)\}_S$ is Cauchy.*

Proof. Since we know that f is uniformly continuous, given $\varepsilon > 0$ there is $\delta > 0$ such that $\|f(t) - f(s)\|_X < \varepsilon$ for $|t-s| < \delta$, $t, s \in [a,b]$.

Let us fix a subdivision S_δ of $[a,b]$ with $\|S_\delta\| < \delta$ and, to simplify the notation, let us assume that S is a refinement of S_δ generated by splitting the first interval of S_δ into two. That is to say, if the intervals in S_δ have end points $a = t_0 < t_1 < \cdots < t_n = b$, then the intervals in S have end points $a = t_0 < z < t_1 < \cdots < t_n = b$. We can write

$$\begin{aligned} R(S_\delta; c_\delta) &= \sum_{I \in S_\delta} |I|\, f(c_\delta(I)) \\ &= \sum_{i=1}^{n} f(\sigma_i)(t_i - t_{i-1}), \end{aligned}$$

where $\sigma_1, \ldots, \sigma_n$ are the values of any tag c_δ associated with S_δ. Likewise,

$$\begin{aligned} R(S;c) &= \sum_{J \in S} |J| \, f(c(J)) \\ &= f(\mu_1)(z - t_0) + f(\mu_2)(t_1 - z) + \sum_{i=2}^{n} f(\lambda_i)(t_i - t_{i-1}), \end{aligned}$$

where $\mu_1, \mu_2, \lambda_2, \ldots, \lambda_n$ are the values of any tag c associated with S. Then,

$$\begin{aligned} R(S;c) - R(S_\delta;c_\delta) &= f(\mu_1)(z - t_0) + f(\mu_2)(t_1 - z) \\ &\quad - f(\sigma_1)(z - t_0 + t_1 - z) \\ &\quad + \sum_{i=2}^{n} (f(\lambda_i) - f(\sigma_i))(t_i - t_{i-1}) \\ &= (f(\mu_1) - f(\sigma_1))(z - t_0) + (f(\mu_2) - f(\sigma_1))(t_1 - z) \\ &\quad + \sum_{i=2}^{n} (f(\lambda_i) - f(\sigma_i))(t_i - t_{i-1}). \end{aligned}$$

Using the uniform continuity of f, we have

$$\begin{aligned} \|R(S;c) - R(S_\delta;c_\delta)\|_X &\le \varepsilon \left[(z - t_0) + (t_1 - z) + \sum_{i=2}^{n} (t_i - t_{i-1}) \right] \\ &= \varepsilon (b - a). \end{aligned}$$

If we assume that S' is another subdivision, also obtained when the first interval of S_δ is splitted into two, then

$$\begin{aligned} R(S';c') - R(S_\delta;c_\delta) &= f(\mu_1')(z' - t_0) + f(\mu_2')(t_1 - z') \\ &\quad - f(\sigma_1)(z' - t_0 + t_1 - z') \\ &\quad + \sum_{i=2}^{n} (f(\lambda_i') - f(\sigma_i))(t_i - t_{i-1}) \\ &= (f(\mu_1') - f(\sigma_1))(z' - t_0) \\ &\quad + (f(\mu_2') - f(\sigma_1))(t_1 - z') \\ &\quad + \sum_{i=2}^{n} (f(\lambda_i') - f(\sigma_i))(t_i - t_{i-1}). \end{aligned}$$

Thus,

$$\begin{aligned}\left\|R\left(S';c'\right)-R\left(S_\delta;c_\delta\right)\right\|_X &\le \varepsilon\left[\left(z'-t_0\right)+\left(t_1-z'\right)+\sum_{i=2}^{n}\left(t_i-t_{i-1}\right)\right]\\ &=\varepsilon\left(b-a\right).\end{aligned}$$

Finally,

$$\left\|R\left(S';c'\right)-R\left(S;c\right)\right\|_X \le 2\varepsilon\left(b-a\right).$$

If S and S' were arbitrary refinements of S_δ, we would need to repeat the calculations above, as many times as necessary. Although the idea is simple, the writing becomes involved, so we will say no more.

This completes the proof. □

Rounding up this brief presentation of the integral, we will state now several properties to be used later. The proof of the first four can be easily adapted from [6, pp. 63-65, Theorem 3.3.2 and Corollary 1]. As for the last property, the proof follows step by step the case of a function with real values (see, for instance, [5, p. 160, Theorem 5.14]), replacing the absolute value with the norm in X.

Proposition 2.6. *Let X be a real Banach space and let $f(t)$ and $g(t)$ be continuous functions from $[a,b]$ to X. Then,*

(i) *For all $\alpha,\beta\in\mathbb{R}$*

$$\int_a^b\left[\alpha f(t)+\beta g(t)\right]dt=\alpha\int_a^b f(t)\,dt+\beta\int_a^b g(t)\,dt.$$

(ii) *If $a<z<b$,*

$$\int_a^b f(t)\,dt=\int_a^z f(t)\,dt+\int_z^b f(t)\,dt.$$

(iii)

$$\left\|\int_a^b f(t)\,dt\right\|_X\le(b-a)\sup_{a\le t\le b}\left\|f(t)\right\|_X.$$

(iv) *If $T\in L(X)$,*

$$T\left(\int_a^b f(t)\,dt\right)=\int_a^b T\left(f(t)\right)dt.$$

(v) *The function*

$$t\to\int_a^t f(s)\,ds$$

is continuous and has a continuous derivative, equal to $f(t)$, for each $t\in[a,b]$.

Property (v) shows that the function $t \to \int_a^t f(s)\, ds$ is an antiderivative of the function $f(t)$. Furthermore, it is still true in the vector case that if g_1 and g_2 are antiderivatives of f, then the difference $g_1 - g_2$ is constant on $[a,b]$ (see [4, p. 160, 8.7.1]).

Our next step in justifying (1.5) as the general solution of (1.4), will be to study the exponential function.

3. The Exponential Function

Given a function $f : [a,b] \to L(X)$, we denote by $f^j(t)$ the composition

$$\underbrace{(f(t)) \circ (f(t)) \circ \cdots \circ (f(t))}_{\text{j times}},$$

in the Banach algebra $L(X)$ (see, for instance, [6, p. 51]). It is understood that $f^j(t) = I$, the identity operator in $L(X)$, when $j = 0$.

The proof of the following lemma uses basic properties of the norm in $L(X)$ and it will be omitted:

Lemma 3.1. *The series* $\sum_{j\geq 0} \frac{f^j(t)}{j!}$ *converges in* $L(X)$ *for each* $t \in [a,b]$.

Definition 3.2. For each $t \in [a,b]$, the exponential $e^{f(t)}$ is the operator in $L(X)$ defined as

$$e^{f(t)} = \sum_{j\geq 0} \frac{f^j(t)}{j!}.$$

Lemma 3.3. *The operator* $e^{f(t)}$ *is invertible in* $L(X)$ *with inverse*

$$e^{-f(t)} = \sum_{j\geq 1} (-1)^j \frac{f^j(t)}{j!}.$$

Proof. We have to prove that

$$e^{-f(t)} \circ e^{f(t)} = e^{f(t)} \circ e^{-f(t)} = I. \tag{3.1}$$

In fact,

$$
\begin{aligned}
e^{-f(t)} \circ e^{f(t)} &= \left(\sum_{j\geq 0} (-1)^j \frac{f^j(t)}{j!}\right) \circ \left(\sum_{l\geq 0} \frac{f^l(t)}{l!}\right) \\
&= \sum_{k\geq 0} \sum_{j+l=k} (-1)^j \frac{\left(f^j(t)\right) \circ \left(f^l(t)\right)}{j!l!} \\
&= \sum_{k\geq 0} \frac{f^k(t)}{k!} \sum_{j=0}^{k} (-1)^j \frac{k!}{j!\,(k-j)!} \\
&= \sum_{k\geq 0} \frac{f^k(t)}{k!} (1+(-1))^k = I,
\end{aligned}
$$

and the proof of the other part of (3.1) is similar.

This completes the proof. □

Other algebraic properties will have expected from the exponential are true as well, sometimes under certain assumptions. For instance, if $f(t)$ and $g(t)$ commute for each $t \in [a,b]$, then

$$e^{f(t)} \circ e^{g(t)} = e^{f(t)+g(t)}.$$

Lemma 3.4. *If the function $f : [a,b] \to L(X)$ is continuous, then the exponential $t \to e^{f(t)}$ is continuous as well, from $[a,b]$ to $L(X)$.*

Proof. For $t, s \in [a,b]$,

$$e^{f(t)} - e^{f(s)} = \sum_{j\geq 1} \frac{f^j(t) - f^j(s)}{j!}.$$

The identity

$$f^j(t) - f^j(s) = \sum_{l=1}^{j} \left(f^{j-l}(t)\right) \circ (f(t) - f(s)) \circ \left(f^{l-1}(s)\right),$$

can be easily verified by performing the operations indicated in the right hand side. Since

$$\left| \|f(t)\|_{L(X)} - \|f(s)\|_{L(X)} \right| \leq \|f(t) - f(s)\|_{L(X)},$$

the real function $t \to \|f(t)\|_{L(X)}$ is continuous, and bounded, on $[a,b]$. Thus, there exists $M > 0$ so that

$$\left\|f^j(t) - f^j(s)\right\|_{L(X)} \leq jM^{j-1} \|f(t) - f(s)\|_{L(X)},$$

and

$$\left\|e^{f(t)} - e^{f(s)}\right\|_{L(X)} \leq e^M \|f(t) - f(s)\|_{L(X)},$$

which shows that the exponential function is continuous, and also uniformly continuous, on $[a,b]$.

This completes the proof. □

Proposition 3.5. *If the function $f:[a,b]\to L(X)$ is differentiable, then the exponential $t\to e^{f(t)}$ is also differentiable from $[a,b]$ to $L(X)$. Moreover,*

$$\frac{d}{dt}e^{f(t)}=\sum_{j\geq 1}\frac{1}{j!}\sum_{l=1}^{j}\left(f^{j-l}(t)\right)\circ\left(f'(t)\right)\circ\left(f^{l-1}(t)\right), \tag{3.2}$$

which reduces to the usual formula,

$$\frac{d}{dt}e^{f(t)}=e^{f(t)}\circ\left(f'(t)\right), \tag{3.3}$$

when $f(t)$ and $f'(t)$ commute. If this is the case, we can also write

$$e^{f(t)}\circ\left(f'(t)\right)=\left(f'(t)\right)\circ e^{f(t)}. \tag{3.4}$$

Proof. If $f(t)$ and $f'(t)$ commute for each $t\in[a,b]$, it should be clear that the right hand side of (3.2) reduces to the right hand side of (3.3). Moreover, the equality (3.4) holds. So, we are left to prove (3.2). We begin by writing

$$\frac{e^{f(s)}-e^{f(t)}}{s-t}=\sum_{j\geq 1}\frac{1}{j!}\sum_{l=1}^{j}\left(f^{j-l}(s)\right)\circ\left(\frac{f(s)-f(t)}{s-t}\right)\circ\left(f^{l-1}(t)\right).$$

Next, we claim that we can take the limit in $L(X)$, as $s\to t$, obtaining

$$\frac{d}{dt}e^{f(t)}=\sum_{j\geq 1}\frac{1}{j!}\sum_{l=1}^{j}\left(f^{j-l}(t)\right)\circ\left(f'(t)\right)\circ\left(f^{l-1}(t)\right).$$

To justify the existence of the limit, we reason as follows:

For a fixed $t\in[a,b]$ and $s\in[a,b]$, we define the function $F_t:[a,b]\to L(X)$ as

$$F_t(s)=\begin{cases}\frac{f(s)-f(t)}{s-t} & \text{if } s\neq t,\\ f'(t) & \text{if } s=t.\end{cases}$$

This function is continuous from $[a,b]$ to $L(X)$. Likewise, if we consider

$$G_{t,j}(s)=\frac{1}{j!}\sum_{l=1}^{j}\left(f^{j-l}(s)\right)\circ\left(F_t(s)\right)\circ\left(f^{l-1}(t)\right),$$

the function $G_{t,j}:[a,b]\to L(X)$ is also continuous.

We claim that the series $\sum_{j\geq 1}G_{t,j}(s)$ converges in $L(X)$, uniformly with respect to $s\in[a,b]$. To prove this claim, we will invoke the vector version of the so called Weierstrass criterion (see, for instance, [5, p. 219, Theorem 7.2]), asserting that the series will converge in $L(X)$, uniformly with respect to s in $[a,b]$, if

$$\sup_{a\le s\le b}\|G_{t,j}(s)\|_{L(X)}\le M_{t,j},$$

and the series $\sum_{j\ge 1} M_{t,j}$ converges in $\mathbb{R}$, for a fixed values of $t\in[a,b]$.

Since the function $F_t(s)$ is continuous, there is $C_t>0$ such that

$$\sup_{a\le s\le b}\|F_t(s)\|_{L(X)}\le C_t.$$

Then,

$$\begin{aligned}\|G_{t,j}(s)\|_{L(X)} &\le \frac{C_t}{j!}\sum_{l=1}^{j}\left\|f^{j-l}(s)\right\|_{L(X)}\left\|f^{l-1}(t)\right\|_{L(X)}\\ &\le \frac{C_t}{(j-1)!}\left(\sup_{a\le s\le b}\|f(s)\|_{L(X)}\right)^{j-1},\end{aligned}$$

so, the numbers $M_{t,j}$ exist. As a consequence, there exists

$$\lim_{s\to t}\sum_{j\ge 1}G_{t,j}(s)=\sum_{j\ge 1}\left(f^{j-l}(t)\right)\circ\left(f'(t)\right)\circ\left(f^{l-1}(t)\right).$$

Finally, since for $s\ne t$ we have the equality

$$\frac{e^{f(s)}-e^{f(t)}}{s-t}=\sum_{j\ge 1}G_{t,j}(s),$$

we must have

$$\lim_{s\to t}\frac{e^{f(s)}-e^{f(t)}}{s-t}=\lim_{s\to t}\sum_{j\ge 1}G_{t,j}(s).$$

That is to say, we get (3.2).

This completes the proof. □

Proposition 3.5 implies that, as we take the derivative of the function $y(t)$ in (1.5), we will be able to use the chain rule, that is to say,

$$\frac{d}{dt}e^{\int u(t)dt}=\left(e^{\int u(t)dt}\right)\circ(u(t)),$$

if $u(t)$ and $\int u(t)\,dt$ commute.

Let us see now a couple of examples illustrating this situation.

Example 3.6. We consider the function $u(t)=\alpha(t)T$, where $\alpha:[a,b]\to\mathbb{R}$ is continuous with a continuous derivative, and $T\in L(X)$. Since

$$\|u(s)-u(t)\|_{L(X)}=|\alpha(s)-\alpha(t)|\,\|T\|_{L(X)},$$

the function $u : [a, b] \to L(X)$ is continuous. Thus, the function

$$w(t) = \int_a^t u(s)\, ds = \left(\int_a^t \alpha(s)\, ds \right) T = \beta(t)\, T,$$

is an antiderivative of u. We claim that $u(t)$ and $w(t)$ commute, for each $t \in [a, b]$. In fact,

$$\begin{aligned} (u(t)) \circ (w(t)) &= (\alpha(t)\, T) \circ (\beta(t)\, T) \\ &= (\alpha(t)\, \beta(t))\, (T \circ T) \\ &= (w(t)) \circ (u(t)). \end{aligned}$$

Let us observe that our example includes the constant coefficient case, when $\alpha(t) = 1$ for all $t \in [a, b]$. In this case, we can choose as integrating factor,

$$\mu(t) = e^{-tu},$$

for $t \in [a, b]$. The function $\mu : [a, b] \to L(X)$ is continuous and it has a derivative. Moreover,

$$\mu'(t) = -(\mu(t)) \circ u, \tag{3.5}$$

so, $\mu'(t)$ is continuous .

As for (1.5), we have to consider $\int_a^t (e^{-su})(v(s))\, ds$. The function to be integrated is continuous from $[a, b]$ to X, so the integral exists. Moreover,

$$\frac{d}{dt} \int_a^t \left(e^{-su}\right)(v(s))\, ds = \left(e^{-tu}\right)(v(t)).$$

Finally, using (iv) in Proposition 2.6, we can conclude that

$$y(t) = \int \left(e^{(t-s)u}\right)(v(s))\, ds + \left(e^{tu}\right)(C), \tag{3.6}$$

is the general solution of (1.4).

In the same way, we can justify the operations in (1.5), any time that $u(t)$ and $\int_a^t u(s)\, ds$ commute.

Our next example will show that they can not always commute.

Example 3.7. We pick a real and separable Hilbert space H, denoting $\{x_j\}_{j \geq 1}$ as an orthonormal basis (see, for instance, [10, Section 5.17, p. 305]). If $t \in [0, 1]$ and

$$x = \sum_{j \geq 1} \alpha_j x_j,$$

we define

$$(u(t))(x) = \sum_{j \geq 1} \frac{\alpha_j}{(t + j)^2} x_{j+1}. \tag{3.7}$$

Since

$$\sum\nolimits_{j\geq 1} \frac{\alpha_j^2}{(t+j)^4} \leq \sum\nolimits_{j\geq 1} \alpha_j^2,$$

then (3.7) defines, for each $t \in [0,1]$, an operator in $L(H)$ with $\|u(t)\|_{L(H)} \leq 1$. This operator is a modification of the left-shift operator T_l (see, for instance, [10, p. 422, Example 5]), defined as

$$T_l\left(\sum_{j\geq 1} \alpha_j x_j\right) = \sum_{j\geq 1} \alpha_j x_{j+1}.$$

We claim that the function $u : [0,1] \to L(H)$ given by (3.7), is continuous. Indeed, if we fix $s, t \in [0,1]$, we can write

$$\begin{aligned}\left|\frac{1}{(s+j)^2} - \frac{1}{(t+j)^2}\right| &= |s-t| \frac{t+s+2j}{(s+j)^2(t+j)^2} \\ &\leq 2\frac{j+1}{j^4}|s-t| \leq 4|s-t|.\end{aligned}$$

So, for a fixed $x \in H$, we have

$$\begin{aligned}\|(u(s))(x) - (u(t))(x)\|_H^2 &= \sum_{j\geq 1} \alpha_j^2 \left(\frac{1}{(s+j)^2} - \frac{1}{(t+j)^2}\right)^2 \\ &\leq 16(s-t)^2 \|x\|_H^2.\end{aligned}$$

That is to say,

$$\begin{aligned}\|u(s) - u(t)\|_{L(H)} &= \sup_{\|x\|_H \leq 1} \|(u(s) - u(t))(x)\|_H \\ &\leq 4|s-t|.\end{aligned}$$

So, the function $u : [0,1] \to L(H)$ is continuous.

Let us consider next

$$(w(t))(x) = -\sum\nolimits_{j\geq 1} \frac{\alpha_j}{t+j} x_{j+1}.$$

Then,

$$\|(w(t))(x)\|_H^2 = \sum_{j\geq 1} \frac{\alpha_j^2}{(t+j)^2} \leq \sum_{j\geq 1} \alpha_j^2 = \|x\|_H^2,$$

which tells us that $w(t) \in L(H)$ for each $t \in [0,1]$. Moreover, the function $w : [0,1] \to L(H)$ is continuous as well, the proof is similar to the way we proved the continuity of $u(t)$. Now, we want to show that $w(t)$ is an antiderivative of $u(t)$. To this purpose, we fix $x \in H$ and write

$$\frac{(w(s))(x) - (w(t))(x)}{s-t} - (u(t))(x)$$

$$= \sum_{j\geq 1} \frac{\alpha_j}{s-t}\left(-\frac{1}{s+j}+\frac{1}{t+j}-\frac{s-t}{(t+j)^2}\right)x_{j+1}.$$

A few algebraic manipulations will show that

$$\frac{1}{s-t}\left(-\frac{1}{s+j}+\frac{1}{t+j}-\frac{s-t}{(t+j)^2}\right)=\frac{t-s}{(s+j)(t+j)^2}.$$

As a consequence,

$$\left\|\frac{(w(s))(x)-(w(t))(x)}{s-t}-(u(t))(x)\right\|_H^2 \leq (t-s)^2\|x\|_H^2.$$

That is to say,

$$\lim_{s\to t}\left\|\frac{w(s)-w(t)}{s-t}-u(t)\right\|_{L(H)}=0.$$

We contend that, for a fixed $t\in[0,1]$, the operators $u(t)$ and $w(t)$ do not commute. In fact,

$$((w(t))\circ(u(t)))(x)=(w(t))\left(\sum_{j\geq 1}\beta_j^t x_j\right),$$

where

$$\beta_j^t=\begin{cases} 0 & j=1 \\ \frac{\alpha_{j-1}}{(t+j-1)^2} & j\geq 2 \end{cases}.$$

So,

$$\begin{aligned}((w(t))\circ(u(t)))(x) &= -\sum_{j\geq 1}\frac{\beta_j^t}{t+j}x_{j+1}\\ &= -\sum_{j\geq 2}\frac{\alpha_{j-1}}{(t+j-1)^2(t+j)}x_{j+1}\\ &= -\sum_{j\geq 1}\frac{\alpha_j}{(t+j)^2(t+j+1)}x_{j+2}.\end{aligned}$$

However, a similar calculation will show that

$$((u(t))\circ(w(t)))(x)=-\sum_{j\geq 1}\frac{\alpha_j}{(t+j)(t+j+1)^2}x_{j+2}.$$

To conclude our exposition, we will discuss some examples.

4. Three Illustrations of the Integrating Factor Method

First, we consider the case $X = \mathbb{R}^n$, the function $u(t)$ being a constant real matrix with n rows and n columns. In this case, (1.4) is a system of n ordinary differential equations with constant coefficients, in n unknowns. As we observed in Example 3.6, the formula (1.5) reduces to (3.6). That is to say,

$$y(t) = \int \left(e^{(t-s)u}\right)(v(s))\,ds + \left(e^{tu}\right)(C), \tag{4.1}$$

for any $C \in \mathbb{R}^n$.

The difficulty with this formula is having to deal with the convergence of the series defining each exponential and then, describing the entries of the resulting matrices. However, if the matrix u is diagonalizable, (4.1) has an explicit form. Let us recall that a matrix u is diagonalizable if (see, for instance, [9, p. 507]) there is a non singular matrix p and a diagonal matrix d such that

$$p^{-1}up = d. \tag{4.2}$$

There are simple conditions that characterize a diagonalizable matrix (see, for instance, [9, p. 512]).

From (4.2), we can write

$$e^{tu} = e^{tpdp^{-1}} = \sum_{j\geq 0} \frac{t^j}{j!}\left(pdp^{-1}\right)^j.$$

Let us observe that

$$\left(pdp^{-1}\right)\left(pdp^{-1}\right) = pdp^{-1}pdp^{-1} = pd^2p^{-1},$$

and inductively,

$$\left(pdp^{-1}\right)^j = pd^jp^{-1},$$

for all $j \geq 1$. Then,

$$e^{tu} = \sum_{j\geq 0} \frac{t^j}{j!} pd^jp^{-1} = pe^{td}p^{-1}.$$

Using its series representation, we can see that e^{td} is a diagonal matrix, which shows that e^{tu} is diagonalizable. Then, (4.1) can be written as

$$y(t) = \int \left(pe^{(t-s)d}p^{-1}\right)(v(s))\,ds + pe^{td}p^{-1}C,$$

for any $C \in \mathbb{R}^n$.

Finally, since d is a diagonal matrix, we do not need to use the series to define e^{td}. Instead, we can use the eigenvalues $\lambda_1, \ldots, \lambda_n$ of the matrix u, counted with their multiplicity. In fact (see, for instance, [9, p. 525]), e^{td} is the diagonal matrix with the values $e^{t\lambda_1}, \ldots, e^{t\lambda_n}$ along the diagonal.

As a consequence, these values turn out to be the eigenvalues of e^{td}. Much can be said about the computational aspects of our calculations, and about extensions to non diagonalizable matrices (see, for instance, [9, pp. 528 and 599]).

In our second illustration, the Banach space X is the Hilbert space $l^2 = l^2(\mathbb{N})$ of those real sequences that are square summable, with the norm associated to its inner product $\langle , \rangle_{l^2}$ (see, for instance, [10, p. 280, Example 9]).

We fix an operator $u \in L(l^2)$ that we assume to be self adjoint and compact (see, for instance, [11, pp. 188 and 190]). The spectral theorem (see, for instance, [11, p. 190, Theorem 6.2]) assures the existence of an orthonormal basis $\{f_n\}_{n\geq 1}$ of l^2 and a sequence of real numbers $\{\lambda_n\}_{n\geq 1}$ converging to zero, such that $u(f_n) = \lambda_n f_n$ for each $n \in \mathbb{N}$. That is to say, the basis consists of the eigenvectors of u associated with its eigenvalues $\{\lambda_n\}_{n\geq 1}$. Since

$$\begin{aligned} \left(e^{tu}\right)(f_n) &= \sum_{j\geq 0} \frac{t^j}{j!} \left(u^j\right)(f_n) \\ &= \sum_{j\geq 0} \frac{t^j \lambda_n^j}{j!} f_n \\ &= \left(e^{t\lambda_n}\right)(f_n), \end{aligned}$$

we can see that, for each $x \in l^2$,

$$x = \sum_{n\geq 1} \alpha_n f_n,$$

we have

$$\left(e^{tu}\right)(x) = \sum_{n\geq 1} \alpha_n e^{t\lambda_n} f_n.$$

The operator e^{tu} is not compact because the sequence $\left\{e^{t\lambda_n}\right\}_{n\geq 1}$ of its eigenvalues, does not converge to zero as $n \to \infty$ (see, for instance, [11, p. 190, Theorem 6.2]).

Given the continuous function $v : [a, b] \to l^2$ in (4.1), the properties of the inner product imply that

$$v(t) = \sum_{n\geq 1} v_n(t) f_n,$$

with $v_n(t) = \langle v(t), f_n\rangle_{l^2}$ is continuous from $[a, b]$ to $\mathbb{R}$. Moreover, if

$$C = \sum_{n\geq 1} c_n f_n$$

for any sequence $\{c_n\}_{n\geq 1} \in l^2$, then we can write (4.1) as

$$y(t) = \sum_{n\geq 1} \left(\int v_n(s) e^{(t-s)\lambda_n} ds \right) f_n + \sum_{n\geq 1} c_n e^{t\lambda_n} f_n.$$

For our third and last illustration, we assume that the real Banach space X has a basis $\{x_n\}_{n\geq 1}$. That is to say, for each $x \in X$ there exists a unique sequence $\{\alpha_n\}_{n\geq 1}$ of real numbers, so that $x = \sum_{n\geq 1} \alpha_n x_n$. As typical examples of this situation, we mention the sequence spaces $l^p = l^p(\mathbb{N})$, for $1 \leq p < \infty$, and the space $C[a,b]$ of the real continuous functions on $[a,b]$ (see, for instance, [7, p. 625]). This reference, [7], gives an excellent, short presentation on the subject of bases in a Banach space.

The uniqueness condition on the coefficients $\{\alpha_n\}_{n\geq 1}$ implies that, for each $n \geq 1$, the functional $l_n(x) = \alpha_n$ is well defined and linear. Moreover (see, for instance, [7, p. 627]), there exists $C > 0$, actually independent of n, so that

$$|l_n(x)| \leq C \|x\|_X, \tag{4.3}$$

showing that l_n is bounded, or continuous, on X, for all $n \geq 1$.

Let us now assume that $u \in L(X)$ is the projection onto a non trivial subspace of X (see, for instance, [10, p. 201]). Then, $u^2 = u$ and, consequently, $u^j = u$ for every $j \geq 1$. Thus,

$$\begin{aligned}\left(e^{tu}\right)(x_n) &= \sum_{j\geq 0} \frac{t^j}{j!} \left(u^j\right)(x_n) \\ &= x_n + \left(e^t - 1\right) u(x_n),\end{aligned}$$

and

$$\begin{aligned}\left(e^{tu}\right)(x) &= \sum_{n\geq 1} \alpha_n \left(x_n + \left(e^t - 1\right) u(x_n)\right) \\ &= \left(I + \left(e^t - 1\right) u\right)(x).\end{aligned}$$

Given the continuous function $v(t)$ in (4.1), the estimate (4.3) implies that

$$v(t) = \sum_{n\geq 1} l_n(v(t)) x_n = \sum_{n\geq 1} v_n(t) x_n,$$

with $v_n : [a,b] \to \mathbb{R}$ is continuous. So, finally, (4.1) can be written as

$$y(t) = \sum_{n\geq 1} \int v_n(s) \left(x_n + \left(e^{t-s} - 1\right) u(x_n)\right) ds + C + \left(e^t - 1\right) u(C),$$

for any $C \in X$.

References

1. A. Banner, *The Calculus Lifesaver*, Princeton University Press, New Jersey, 2007.
2. G. Birkhoff, *Moore-Smith Convergence in General Topology*, Ann. of Math., 38 (1937), pp. 39–56.
3. P. Blanchard, R.L. Devaney, and G. Hall, *Differential Equations*, Brooks/Cole, California, 2006.
4. J. Dieudonné, *Foundations of Modern Analysis*, Academic Press, New York, 1960.
5. E.D. Gaughan, *Introduction to Analysis*, Brooks/Cole, California, 1998.
6. E. Hille and R.S. Phillips, *Functional Analysis and Semi-Groups*, American Mathematical Society, Rhode Island, 1957.
7. R.C. James, *Bases in Banach spaces*, Amer. Math. Monthly, 89 (1982), pp. 625–640.
8. J.L. Kelley, *General Topology*, D. Van Nostrand Co., New Jersey, 1955.
9. C.D. Meyer, *Matrix Analysis and Applied Linear Algebra*, Society for Industrial and Applied Mathematics, Pennsylvania, 2000.
10. A.W. Naylor and G.R. Sell, *Linear Operator Theory in Engineering and Science*, Springer-Verlag, New York, 1982.
11. E.M. Stein and R. Shakarchi, *Real Analysis: Measure Theory, Integration and Hilbert Spaces*, Princeton University Press, New Jersey, 2005.

---

[1] Department of Mathematics, New Mexico State University, Las Cruces, New Mexico 88003, USA.

E-mail address: `jalvarez@nmsu.edu`

[2] Departamento de Matemáticas, Universidad de Sonora, Hermosillo, Sonora 83000, Mexico.

E-mail address: `carolina.espinoza@mat.uson.mx`

[3] Departamento de Matemáticas, Universidad de Sonora, Hermosillo, Sonora 83000, Mexico.

E-mail address: `martha@mat.uson.mx`