Strategy for the teaching of linear differential equations with constant coefficients through the inverse operator method

J Aréchiga Maravillas¹, J Mejía Brito¹ and S Arceo Díaz¹

¹ Instituto Tecnológico de Colima, Tecnológico Nacional de México, Colima, México

E-mail: santiago.arceo@itcolima.edu.mx

Abstract. One of the most desirable aspects of teaching mathematics is the development of a wide variety of strategies that improve the quality of learning in students. This work shows how the implementation of a method based on inverse operators to solve linear differential equations produces better results than other, more traditional, methods. We tested our hypothesis over a group conforming by 80 students from the Biochemistry and Mechatronics Engineering Departments of the Technological Institute of Colima, part of the National Technological Institute of México. The experimental design was comprised by two groups of students: the traditional methods were implemented to the first group of students (named as the control group) while the inverse operator method was taught to the other (named as the test group). The statistical test over the results shows a significant difference between the approval rating of both groups, suggesting that there is enough evidence to conclude that the teaching of inverse operators is an efficient alternative for the solution of linear differential equations with constant coefficients.

1. Introduction

A differential equation is, according to Varona [1], an equation relating a function, and its independent variable, to its derivative. If the equation contains only derivatives of the function to a single independent variable, it is defined as an ordinary differential equation (ODE), while if it contains derivatives to multiple independent variables it is called a partial differential equation (PDE) [2,3].

Learning how to solve differential equations is an important skill for anyone interested in the mathematical study and modeling of natural phenomena [4–6]. According to González [7] there are many areas of knowledge in which being able to solve differential equations allows a better understanding of natural phenomena.

The development of new techniques that could help the student not only to achieve mathematical proficiency but also to gain a better understanding of the processes involved in solving differential equations is of great importance [8,9]. Whitehead [10] states that “One of the worst problems in the development of mathematical skills is to be able to explain to students the purpose behind each technique”. Learning a variety of methods for solving differential equations is essential for engineering students [11,12].

Within the National Technological Institute of México (Tecnológico Nacional de México), the educative model is based on the development of a rich set of different competencies in students.
Specifically, the subject of differential equations is considered important since it helps to gain experience in the creation of mathematical models that can be used to describe a wide variety of phenomena related to multiple sciences as physics, chemistry or biology. The second unit, within the program of study for differential equations, aims to teach the student how to solve linear differential equations. However, operator methods, like the one described in this work, are not considered and all the teaching is made from a more traditional point of view, as the undetermined coefficient and parameter variation methods.

Operator methods have various advantages: they allow a better comprehension on the concepts involved in process of solving linear differential equations, they provide a way to relate the subject of differential equations to linear algebra and basic calculus and their simplicity help students to perform much better, as is discussed in the view of evidence on the last section.

2. Theory
2.1. Linear operators
It is said that a linear transformation \( L : C^n(I) \rightarrow C(I) \) is an nth-order linear differential operator over the interval \( I \), if it can be written as:
\[
L = a_n(x)D^n + a_{n-1}(x)D^{n-1} + ... + a_1(x)D + a_0(x),
\]
(1)
where the coefficients \( a_0(x),...,a_n(x) \) are continuous functions through and fulfill the condition that the leading coefficient, \( a_n(x) \), is different from zero. The image of \( f(x) \) in \( C(I) \), under a linear differential operator is defined by
\[
Lf(x) = a_n(x)\frac{d^n}{dx^n}f(x) + ... + a_1(x)\frac{d}{dx}f(x) + a_0(x)f(x),
\]
(2)
or in a more compact way:
\[
Ly = a_n(x)y^{(n)} + ... + a_1(x)y' + a_0(x)y,
\]
(3)
where \( y, ..., y^{(n)} \) are all the consecutive derivatives of \( f(x) \).

2.2. Linear differential equations
An nth-order linear differential equation, defined within the interval \( I \), is an equation which follows the operator form:
\[
Ly = h(x),
\]
(4)
where \( h(x) \) is continuous within \( I \), and \( L \) is a linear differential operator of nth-order, well defined within that interval. The equation is said to be homogeneous if \( h(x) \) is identically zero through \( I \). Otherwise, it is called a non-homogeneous equation. The function \( y(x) \) is a solution to a differential equation if, and only if, \( y(x) \) belongs to \( C^n(I) \) and identically satisfies the equation through \( I \).

2.3. Linear differential equations as linear transformations
From the point of view of linear algebra, a linear differential equation can be conceptualized as a linear transformation. In the case of the nth-order linear homogeneous differential equation
\[
Ly = 0.
\]
(5)
If \( Ly = 0 \) is assumed to be a linear transformation, then looking for its solution would be equivalent to searching for the null space of \( L \), that would give the solution space. Given that the domain of \( C^n(I) \) is a vector space, it is assumed that the solution space of the nth-order
homogeneous differential equation is an n-dimensional subspace of $C^n(I)$, so it is necessary to find a base comprised of n vectors, n linearly independent solutions $y_1(x), ..., y_n(x)$ to form that base. So, any solution can be described as a linear combination of the vectors within the base $y(x) = c_1y_1(x) + c_2y_2(x) + ... + c_ny_n(x)$.

The linear transformation view can also be applied to non-homogeneous linear transformation: if $yp$ is a particular solution for the equation $Ly = h(x)$ and if $yh$ is the general solution for the associated homogeneous equation $Ly = 0$, then $yp + yh$ is the general solution for $Ly = h(x)$. In other words, the solution set of the non-homogeneous linear differential equation can be found by summing up all the solutions of the homogeneous equation to any particular solution of the given differential equation.

2.4. The derivative as an operator
The first step is to find a base for the for the nucleus of the operator that will represent the solution space of the differential equation:

$$\{y(x) : y(x) = y(x) = c_1y_1(x) + ... + c_ny_n(x)\},$$

(6)

where $c_1, c_1, ..., c_n$ are real coefficients.

To find a particular solution for a non-homogeneous differential equation we will employ the concept of an inverse operator. To warrant the existence of its inverse transformation, the domain of the transformation must be defined.

Let $L : H → C(I)$ a one to one transformation, then its inverse transformation is given by $L^{-1} : C(I) → H$ where $H = \{y : yC^n(I)\}$ does not contains vectors generated by the basis of the kernel.

In this way if $L(y) = w(x)$, then $L^{-1}(w) = y(x)$, with $y(x)$ as the unique solution and that does not contain vectors generated by the basis of the nucleus. Here is necessary to prove the theorem:

Let $L(D) = a_nD^n + a_{n-1}D^{n-1} + ... + a_1D + a_0$ a linear transformation such as $L : C^n(I) → Nu(L) → C(I)$, then $L$ is a one to one linear transformation.

The demonstration can be done if one proves that:

$$\forall y_1, y_2 \in \{C^n(I) - Nu(L)\} \ni y_1 \neq y_2 \rightarrow L(D)(y_1) \neq L(D)(y_2).$$

(7)

Lets suppose that

$$\forall y_1, y_2 \in \{C^n(I) - Nu(L)\} \ni y_1 \neq y_2 \land L(D)(y_1) = L(D)(y_2)$$

$$L(D)(y_1) = L(D)(y_2) \rightarrow L(D)(y_1 - y_2) = 0.$$  

(8)

Then

$$y_1 - y_2 = 0 \lor y_1 - y_2 \in Nu(L) \rightarrow y_1 = y_2 \lor y_1, y_2 \in Nu(L),$$

(9)

what would contradict our initial supposition. Hence:

$$L : C^n(I) - Nu(L) → C(I),$$

(10)

thus proving our theorem and ensuring that the solution found by calculating the inverse operator for $Ly = h(x)$ is unique. The disadvantage of conceptualizing a differential equation as a linear transformation and then finding the solution by searching for the inverse operator is the lack of a general method, meaning that it divides in different cases that depend on the specific form of the function involved in the differential equation.

Let $x$ be the independent variable. One defines the derivative operators:
2.4.3. The derivative operator applied to functions of the form \( \sin(ax) \). Applying the derivative operator results in:

\[
D = \frac{d}{dx}, \quad D^2 = \frac{d^2}{dx^2}, \ldots, D^n = \frac{d^n}{dx^n}
\]

2.4.1. The derivative operator applied to exponential functions. Let \( f(x) = e^{\pm ax} \), where \( a \) is a constant. Applying the derivative operator results in:

\[
D(e^{\pm ax}) = \pm ae^{\pm ax}, \quad D^2(e^{\pm ax}) = \pm a^2 e^{\pm ax}, \ldots, D^n(e^{\pm ax}) = \pm a^n e^{\pm ax}.
\]

Adding all these operators one obtains:

\[
D(e^{\pm ax}) + D^2(e^{\pm ax}) + \ldots + D^n(e^{\pm ax}) = \pm a e^{\pm ax} + (\pm a^2 e^{\pm ax}) + \ldots + (\pm a^n e^{\pm ax}).
\]

Rearranging the terms:

\[
(D + D^2 + D^3 + \ldots + D^n)e^{\pm ax} = ((\pm a) + (\pm a^2) + \ldots + (\pm a^n))e^{\pm ax},
\]

which can be written as:

\[
P(D)e^{\pm ax} = P(\pm a)e^{\pm ax}.
\]

Multiplying by \( \frac{1}{P(D)P(\pm a)} \), one has \( \frac{1}{P(D)}e^{\pm ax} = \frac{1}{P(\pm a)}e^{\pm ax} \), in which \( \frac{1}{P(D)} \) represents the inverse operator for the \( P(D) \) polynomial.

2.4.2. The derivative operator applied to polynomials. We begin with a polynomial function \( a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n \). By applying the derivative operator, one obtains:

\[
L(D)y = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n \rightarrow y_p = \frac{1}{L(D)}(a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n),
\]

where \( \frac{1}{L(D)} \) is a polynomial within \( D \), calculated in a series of nth-order. For example, if \( L(D) = (a_0 + a_1 D + a_2 D^2) \), then

\[
\frac{1}{L(D)} = \frac{1}{a_0} - \frac{a_1}{a_0}D + \frac{-a_2}{a_0}D^2 + \frac{a_3}{a_0}D^3 + \ldots
\]

which would be the third-order series corresponding to a second-order polynomial.

2.4.3. The derivative operator applied to functions of the form \( \sin(ax+b) \), \( \cos(ax+b) \). We begin with functions with the form \( \sin(ax+b) \) and \( \cos(ax+b) \). Applying the derivative operator multiple times, one gets:

\[
D(\sin(ax+b)) = \frac{d}{dx}(\sin(ax+b)) = a \cos(ax+b)
\]

\[
D^2(\sin(ax+b)) = \frac{d^2}{dx^2}(\sin(ax+b)) = -a^2 \sin(ax+b)
\]

\[
D^3(\sin(ax+b)) = \frac{d^3}{dx^3}(\sin(ax+b)) = -a^3 \cos(ax+b)
\]

\[
D^4(\sin(ax+b)) = \frac{d^4}{dx^4}(\sin(ax+b)) = -a^4 \sin(ax+b).
\]
One can see that the sine function repeats for the even order derivatives and this allows for the substitution $D^2 = -a^2$, in a way that $D^4 = (D^2)^2 = (-a^2)^2 = a^4$. From this, it is possible to infer that:

$$y_p = \frac{1}{F(D^2)} \sin (ax + b) = \frac{1}{F(-a^2)} \sin (ax + b)$$  \hspace{1cm} (22)

and for the cosine function:

$$y_p = \frac{1}{F(D^2)} \cos (ax + b) = \frac{1}{F(-a^2)} \cos (ax + b).$$  \hspace{1cm} (23)

2.4.4. Derivative operator applied the multiplication between an arbitrary function and the exponential function. We consider an arbitrary function $u(x)$ that multiplies an exponential function:

$$u(x)e^{\pm ax}$$

We have that

$$P(D)u(x)e^{\pm ax} = e^{\pm ax}P(D \pm a)u(x)$$  \hspace{1cm} (24)

$$y_p = \frac{1}{P(D)} u(x)e^{\pm ax} = \frac{1}{P(D \pm a)} u(x)$$  \hspace{1cm} (25)

2.4.5. multiplication between and arbitrary function and a polynomial. This case is limited to functions of the form $f(x) = x^r u(x)$. Here the solution is found by employing the equation:

$$\frac{1}{P(D)} x^r u(x) = \left[ \left( x + \frac{d}{dD} \right)^r \frac{1}{P(D)} \right] u(x)$$  \hspace{1cm} (26)

3. Results
We tested our working hypothesis by selecting two groups of students taking the differential equations course. The test group, that to which the inverse operator would be taught as an alternative to solve linear differential equations, was comprised by 43 Biochemistry students whose average score ($\mu_2$) was compared afterward against the average score ($\mu_1$) of the control group, conformed by 37 Mechatronics students taking the same course. Our working hypothesis was: The inverse operator method induces a significant improvement on the average score of students learning how to solve linear differential equations against the null hypothesis, stating the opposite. ($H_1 = \mu_2 - \mu_1 > 0$, $H_0 = \mu_2 - \mu_1 \leq 0$).

Given that the number of students was larger than 30, we used the statistical Z-score to test if the difference between the average scores of both groups is significant:

$$Z = \frac{(X_2 - X_1) - (\mu_2 - \mu_1)}{\sqrt{s_2^2/n_2 + s_1^2/n_1}}^{1/2}$$ \hspace{1cm} (27)

We obtained and compared the experimental value of the Z-score against the critical value need to reject the null hypothesis, assuming a confidence level $\alpha = 0.05$: $Z_c = 1.645$. By substituting the results shown in Table 1, we calculate

$$Z = \frac{(78.33 - 74.08) - (0)}{\sqrt{\frac{150.7}{43} + \frac{308.53}{37}}}^{1/2} = 1.23.$$ \hspace{1cm} (28)

Since $Z < Z_c$, we can conclude that our experiment does not give enough evidence to support the claim that reaching students to solve linear differential equations by the inverse operator

5
method does not lead to a significant difference on the average score of the test group, when compared to the average score in the control group.

Table 1. Statistical parameters obtained after the same exam was applied to the test ability of students to solve linear differential equations by employing the inverse operator method (Experimental group) and the traditional method (Control group).

|                  | Experimental group | Control group |
|------------------|--------------------|---------------|
|                  | (43 students)      | (37 students) |
| Average          | 78.3               | 74.1          |
| Median           | 80                 | 76            |
| Mode             | 70                 | 50            |
| Standard dev.    | 12.27              | 17.56         |
| Variance         | 150.7              | 308.5         |
| Percentage of success | 93%         | 73%          |
| Kurtosis         | 0.17               | 1.22          |
| Asymmetry Coefficient | 0.42           | 0.15          |

Afterwards, we searched for evidence on a significant difference on the proportion of students that did not fail the exam about solving linear differential equations. In this case, our working hypothesis is that The percentage of approved students within the test group is significantly larger than for the control group, with the null hypothesis affirming the other possibility \( H_0 : p_2 \leq p_1, H_1 : p_2 > p_1 \). In this case, the Z-score is defined by the equation:

\[
Z = \frac{(\bar{p}_2 - \bar{p}_1) - (p_2 - p_1)}{\left(\frac{(\bar{p}_2)(1-\bar{p}_2)}{n_2} + \frac{(\bar{p}_1)(1-\bar{p}_1)}{n_1}\right)^{1/2}} = \frac{(0.93 - 0.73) - (0)}{(0.03)(0.73) + (0.73)(0.27)}^{1/2} = 2.48. \quad (29)
\]

This value is larger than the corresponding critical value \( Z = 1.645 \), allowing us to conclude that the difference in proportion test gives enough evidence to conclude that the percentage of success in students within the test group is significantly larger than for the control group.

4. Conclusions and future work

With this first study we can remark the following results:

That the inverse operator method induces students to learn faster how to solve linear differential equations, allowing to dedicate more time to the remaining topics on the differential equations course.

Even when the standard Z-test did not give enough evidence to support the hypothesis that employing the inverse operator method leads to a significant improvement on the average rating of the test group, as can be seen in Table 1, the average score for the test group is still higher than that of the control group.

On the other side, there is enough evidence to suggest that the inverse operator method has a significant impact on enhancing the success rate of students when searching for a significant difference on the success rate between two groups with respect to the results obtained for the control group.

We plan to expand our study next semester by repeating this experiment to groups of students from the other disciplines on the Technological Institute of Colima, searching to measure other aspects, as the possible career bias, and to obtain more evidence either in favour or against the teaching the inverse operator method as an alternative of solving linear differential equations.
References
[1] Varona J L 1996 Métodos clásicos de resolución de ecuaciones diferenciales ordinarias (España: Universidad de la Rioja)
[2] Kreider D L, Kuller R G, Ostberg D R 1973 Ecuaciones diferenciales (Bogotá: Fondo Educativo Interamericano)
[3] Zill D G 2002 Ecuaciones diferenciales con aplicaciones en modelado (México: Thompson Learning)
[4] Baranenko G 1988 Problemas y ejercicios de análisis matemático (Moscu: Editorial Mir)
[5] Boyce W E, Diprima R C 2004 Ecuaciones diferenciales y problemas con valores en la frontera (México: Limusa Wiley Editorial)
[6] Nagle R K, Saff E B, Snider A D 2001 Ecuaciones diferenciales y problemas con valores en la frontera (USA: Addison-Wesley)
[7] González T 2000 Revista de Investigación Educativa 18(1) 175-199.
[8] Grossman S I 2007 Álgebra lineal (México: McGraw Hill)
[9] Madalena D M 2009 Enseñanza y aprendizaje en ecuaciones diferenciales con abordaje gráfico numérico y analítico (Burgos: Universidad de Burgos)
[10] Whitehead A N 1965 Journal of the Association of Teachers of Mathematics for the Southeastern Part of England 1(1) 1911
[11] De las fuentes M, Arcos J, Navarro Carlos R 2010 Formación Universitaria 3 33-44
[12] Perdomo J 2011 Revista Didáctica de Matemáticas 78 113-134