CODEE Journal

Volume 8 Article 5

2-7-2011

The Laplace Transform: Motivating the Definition

Howard Dwyer

The Laplace Transform: Motivating the Definition

Howard I. Dwyer
Monmouth College

Keywords: Laplace transform
Manuscript received on June 7, 2010; published on February 7, 2011.

Abstract: Most undergraduate texts in ordinary differential equations (ODE) contain a chapter covering the Laplace transform which begins with the definition of the transform, followed by a sequence of theorems which establish the properties of the transform, followed by a number of examples. Many students accept the transform as a Gift From The Gods, but the better students will wonder how anyone could possibly have discovered/developed it. This article outlines a presentation, which offers a plausible (hopefully) progression of thoughts, which leads to integral transforms in general, and the Laplace transform in particular.

1 Introduction

The most common presentation of the Laplace transform in undergraduate texts on ordinary differential equations (ODE) consists of a definition of the transform, followed by a sequence of theorems which establish the basic properties of the transform, followed by examples in which the Laplace transform is used to solve various types of initial value problems (IVP). The goal of this paper is to outline an alternate presentation, which begins with a development of the concept of an integral transform in general, and follows this with a construction, which "discovers" the Laplace transform. Here are the essential steps in the presentation.

- Discussion of *transformation* as a problem solving technique.
- Review of the concept of a *vector space*, using spaces $\mathbb{R}^2$ and $\mathbb{R}^3$ as familiar examples and introducing a space of functions.
- Review of the *inner product* (dot product) in spaces $\mathbb{R}^2$ and $\mathbb{R}^3$ and definition of an inner product for a space of functions.
- Definition of the *scalar projection* of one vector in the direction of another.
- An example, which shows that information about projections, can be used to find the location of a point. This motivates the idea that we might find it useful to

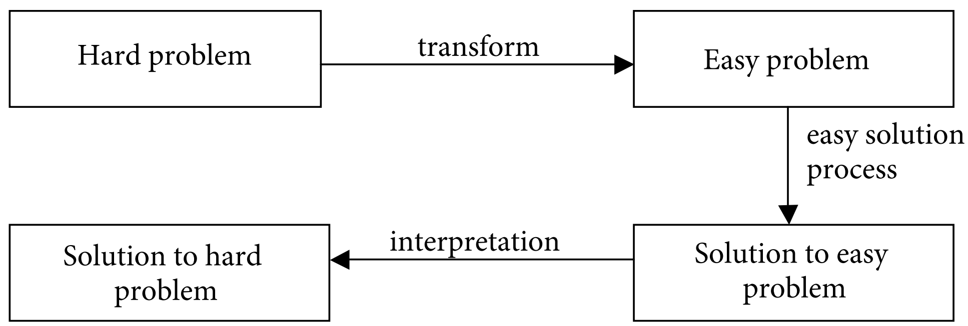
consider the projection of a vector in the directions of multiple vectors, and in particular in the directions defined by a one-parameter family of vectors.

- Definition of the *integral transform*, which can be viewed as vector projection of one function (vector) in the directions of a one-parameter family of functions (vectors).
- A brief discussion of the question of an *inverse for an integral transform.*
- A "wish list" for properties of a transformation to be used as a tool for solving initial value problems.
- Construction of a transform based on the "wish list;" discovery of the *Laplace transform.*

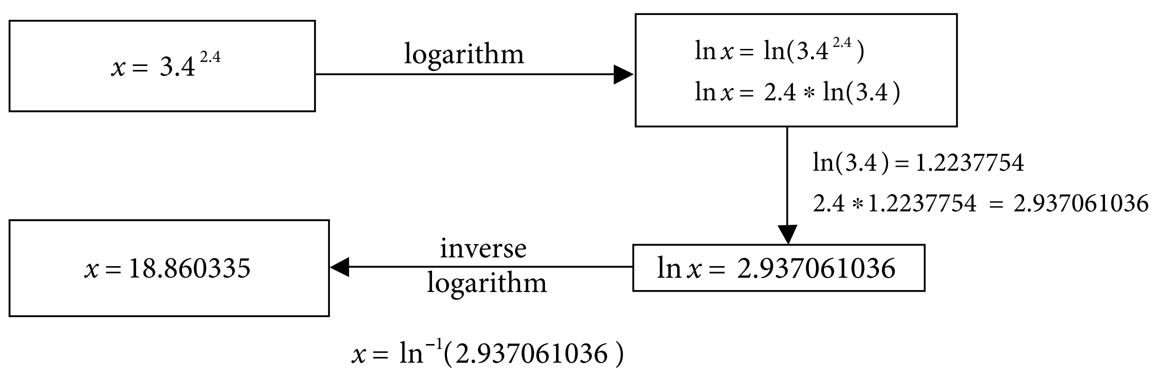
One might look at this outline and wonder how many months of class time this will require, particularly if linear algebra is not a prerequisite for the ODE course. The ODE course at Monmouth College does not have a prerequisite of linear algebra. Students in the course are typically a mixture of science (mostly Physics) and mathematics majors (mostly Secondary Education). Experience indicates that this approach to the transform requires about two additional class periods, compared to the "magic rabbit from a hat" presentation. It is not more than this because it is no longer necessary to develop the properties of the transform after the presentation of its definition; the transform has these properties because it was designed to. The alternative approach also lays the groundwork for a discussion of the inverse—a topic, which sometimes does not get the attention it, deserves.

2 Transformation

We begin the presentation by considering a simple computational problem: compute the value of $x = 3.4^{2.4}$. To get the exact answer, we can use the idea of a *transformation*; we will convert this problem (which is hard) to an *equivalent* problem, which is easy. We will say that two problems are *equivalent* if the solution to one of the problems leads in a straightforward manner to the solution of the other. Remind the students that they have seen this before: as an example, the solution of a system of equations by Gaussian elimination works by transforming the original system into an equivalent system with an easy solution. The concept of solution through transformation can be illustrated with a simple diagram below.



Returning to the computational problem, we show how the logarithm can be used as a transformational tool (on the next page).



What kind of transformation might we use with ODEs? Based on our experience with logarithms, the "dream" would be a transformation, which allows us to replace the operation of differentiation by some easier operation, perhaps something similar to multiplication. Even if we don't get this exactly, coming close might still be useful. We note that the set of differentiable functions forms a vector space; so one possibility is to look at various types of linear transformations. The next stage of the presentation has the goal of convincing students that a reasonable and useful transformation can be defined using an integral and a kernel function with one parameter.

Some courses in ODE will have a prerequisite of linear algebra and others will not; in either case, experience shows that it will be beneficial to review some basic concepts from linear algebra, particularly as they apply to the function spaces consisting of functions f, which are differentiable on some interval $[a, b]$:

- Vector addition and multiplication by a scalar. Closure under addition and scalar multiplication.

  We know how to add two functions, and we know that the result is also going to be a differentiable function. Similarly, we know that multiplying a differentiable function by a constant produces another differentiable function.

- Definition of an inner product.

  In the familiar Euclidean vector spaces $\mathbb{R}^2$ and $\mathbb{R}^3$, we define an inner product (dot product) of two vectors by taking the sum of the products of values from the two vectors. The products are formed using values from the vectors at corresponding places—that is, using corresponding components. Note that in these spaces, the components are referenced by an index which assumes discrete values.

$$\langle \vec{w}, \vec{v} \rangle = \sum_i w_i v_i$$

A function $f(x)$ may be thought of as a vector which has infinitely many components, with the variable x serving as the index for the components. The product of the components from vector f and vector g, at corresponding places within each vector, is therefore $f(x)g(x)$. We take the sum of these products using an integral, allowing the variable x to range over the interval $[a, b]$.

$$\langle f, g \rangle = \int_{x=a}^{x=b} f(x)g(x)\, dx$$

It is important to spend enough time on this idea so that students really see that this definition of an inner product is a natural extension of the notion of dot product with which they are already familiar. If they do not see this, then they will not have any faith in what follows.

- Vector projection.

  The scalar projection of a vector $\vec{v}$ along the direction of $\vec{w}$ is defined in Euclidean spaces by

  $$\text{proj}_{\vec{w}}(\vec{v}) = \frac{\langle \vec{w}, \vec{v} \rangle}{\langle \vec{w}, \vec{w} \rangle}.$$

  In function spaces, the projection of function f along the direction of function g is defined similarly as

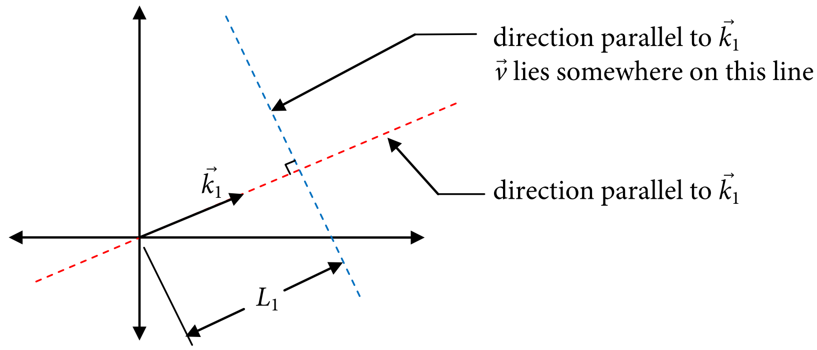
  $$\text{proj}_g(f) = \frac{\langle g, f \rangle}{\langle g, g \rangle}.$$

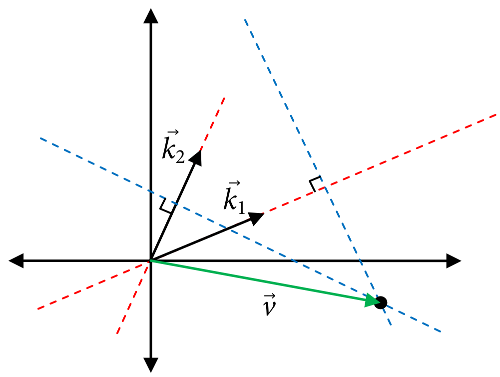
  In either case, we say that two vectors/functions are *orthogonal* if the projection (and hence the inner product) is zero. In Euclidean spaces there is an easy geometric interpretation of orthogonality, in terms of right angles. We should caution students against trying to associate a similar geometric interpretation with orthogonal functions.

3 The "Radar Screen" Analogy

Pose this problem to the students: In $\mathbb{R}^2$, the components of vector $\vec{v}$ are unknown. However, the projected length of $\vec{v}$ onto two independent vectors $\vec{k}_1$ and $\vec{k}_2$ is known. Can we find $\vec{v}$ using the projected lengths?

We can draw a quick sketch to show that if we know the length L_1 of the projection along a known direction $\vec{k}_1$ then we know that the point/vector $\vec{v}$ lies somewhere along a line orthogonal to $\vec{k}_1$.





The projection along a second direction provides another line along which $\vec{v}$ must lie, so we know that $\vec{v}$ lies at the intersection of the two lines.

It is unnecessary to actually do the geometric construction, although students seem to find the visual presentation helpful. We can find the components of the vector $\vec{v}$ quite easily using ordinary algebra.

Example: Suppose that:

1. The inner product of $\vec{v}$ with $\vec{k}_1 = (3, 1)$ is 2.
2. The inner product of $\vec{v}$ with $\vec{k}_2 = (1, 4)$ is 0.5.

Then,

$$\begin{aligned} & \langle \vec{k}_1, \vec{v} \rangle = 3v_x + 1v_y = 2 \\ \text{and} \qquad & \langle \vec{k}_2, \vec{v} \rangle = 1v_x + 4v_y = 0.5. \end{aligned}$$

This system has a unique solution: $\vec{v} = \left(\frac{15}{22}, -\frac{1}{22}\right)$.

A “natural” choice for the direction vectors is to use unit vectors. In $\mathbb{R}^2$, the unit vectors can be characterized using a single parameter

$$\vec{k}(s) = (\cos(s), \sin(s)).$$

As the parameter s increases, we can visualize the direction vector rotating counterclockwise. We can visualize this as being similar (but not quite the same) to the radar screens we see in action movies, with a glowing line rotating around the screen and revealing the location of the target.

What this example illustrates is that it may be useful to consider the inner product of a vector $\vec{v}$ with a collection of other vectors. In this example, we used a one-parameter family of vectors. Using this family of vectors, we have defined a family of functions which take vectors as inputs and return real numbers as results.

Exactly the same procedure can be employed in the space of ordered triples, $\mathbb{R}^3$, although visualizing the procedure geometrically is more challenging. We will need three vectors $\vec{k_1}, \vec{k_2}$, and $\vec{k_3}$ which are linearly independent. Once again, we can define a one-parameter family of unit vectors:

$$\vec{k}(s) = \sqrt{\frac{1}{2}}(1, \cos(s), \sin(s)).$$

This family contains three independent vectors (take $s = 0, s = 1$, and $s = 2$ for example) and thus we can define a one-parameter family of linear functions, such that it is possible to recover the original vector. This example shows that we may not need to use more than one parameter, even in higher dimensions.

4 The "Radar Screen" Analogy Applied to Function Spaces

When the idea of defining a family of linear functions using an inner product with a one-parameter family is applied to function spaces we arrive at the *integral transform*. The one-parameter family is called the *kernel* of the transform. It is common to use a lower-case letter for the name of the function, and the same letter in upper-case for the transform:

$$F(s) = T(f) = \langle k(s), f \rangle = \int_{x=a}^{x=b} k(s,x) f(x) dx.$$

We can define a transform by choosing an interval of integration and a kernel. We should be sure to emphasize at this stage in the presentation that it is not immediately obvious that such transforms are in any way useful, but we will see that they can be. It is also not immediately clear when they must be one-to-one, or what the range might be, or whether it will be possible to "reverse" the process, although the analogy makes this plausible.

5 Constructing a Tool for IVPs

We begin by writing down a characteristic property that the "ideal" transform should have. This is, perhaps, the weakest point in the presentation; there is a "magic rabbit from a hat" nature to this. Experience shows that students do not seem to notice or mind. We have seen that the logarithm turns the operation of exponentiation into the operation of multiplication. Perhaps we can construct a transform which will turn the operation of differentiation into a multiplication. That is, the transform of the derivative would be equal to the transform of the function multiplied by something. One simple choice would be

$$T\left(\frac{d}{dx}f\right) = s \cdot T(f).$$

We will set this as our goal, and see if we can construct an integral transform which has this property, or at least some similar property.

Since IVPs are our interest, we will (quite naturally) look at functions which are defined on the interval $I = \{t | t \geq 0\}$. These functions are solutions to ODE, which means that the functions must be continuous and differentiable (at least once, perhaps more depending on the degree of the ODE). With this choice of interval our integral transform will involve an improper integral. As a necessary condition for the existence of this integral, we must assume that the product of the kernel and the function goes to zero as t gets large, for at least some values of the parameter s,

$$\lim_{t\to\infty}(k(s,t)f(t)) = 0.$$

Note that if any two functions satisfy these restrictions, then every linear combination of them will also, so we are actually considering a subspace of functions.

We now show that we can find a kernel function which (almost) satisfies the goal. Using an integral transform on the interval $I = \{t | t \geq 0\}$, we have

$$T\left(\frac{d}{dx}f\right) = s \cdot T(f),$$

or equivalently,

$$\int_{t=0}^{t=\infty} k(s,t)\left(\frac{df}{dt}\right) dt = s \cdot \int_{t=0}^{t=\infty} k(s,t)(f)\, dt.$$

Since the parameter s is independent of t,

$$\int_{t=0}^{t=\infty} k(s,t)\left(\frac{df}{dt}\right) dt = \int_{t=0}^{t=\infty} s \cdot k(s,t)(f)\, dt.$$

Integration by parts on the first of these integrals yields

$$\int_{t=0}^{t=\infty} k(s,t)\left(\frac{df}{dt}\right) dt = (k(s,t)f(t))_{t=0}^{t=\infty} - \int_{t=0}^{t=\infty} \left(\frac{\partial}{\partial t}k(s,t)\right)(f)\, dt.$$

Since $\lim_{t\to\infty}(k(s,t)f(t)) = 0$, integration by parts simplifies to

$$\int_{t=0}^{t=\infty} k(s,t)\left(\frac{df}{dt}\right) dt = -k(s,0)f(0) + \int_{t=0}^{t=\infty} -\left(\frac{\partial}{\partial t}k(s,t)\right)(f)\, dt.$$

Ignoring (for the moment) the term $k(s,0)f(0)$ and comparing this result to our goal, we see that we will have success if

$$\int_{t=0}^{t=\infty} s \cdot k(s,t)\,(f) dt = \int_{t=0}^{t=\infty} -\left(\frac{\partial}{\partial t}k(s,t)\right)(f)\, dt.$$

This may be easily achieved by setting the integrands equal:

$$s \cdot k(s,t)(f) - \left(\frac{\partial}{\partial t}k(s,t)\right)(f)$$
$$s \cdot k(s,t) = -\left(\frac{\partial}{\partial t}k(s,t)\right)$$
$$\frac{\partial}{\partial t}k(s,t) = -s \cdot k(s,t).$$

Treating the parameter s as a constant (since its value is chosen independently of the value of t) we have a simple, first order differential equation. Students are excited at this point, since this is an equation with which they have extensive experience from the earlier part of the ODE course. Any kernel which satisfies this equation will define an integral transform with (almost) the desired property! And the winner is?

$$k(s,t) = e^{-st}$$

Definition: The *Laplace Transform* of a function $f(t)$ is

$$F(s) = \mathscr{L}(f) = \int_{t=0}^{t=\infty} e^{-st} f(t) dt.$$

We neglected a term in this process, and now must go back and see what property the transform actually has. We see that the Laplace transform has the following property.

Theorem: $\mathscr{L}\left(\frac{d}{dt}f\right) = s\mathscr{L}(f) - f(0)$

In fact, we see that it has this property *because it was designed to.* It's not magic, nor extraordinary good luck, nor trail-and-error, but rather a logical process which produces the Laplace transform as a tool.

6 Conclusion

It is still a magic show, in some ways, but has been well received when tested on students in the introductory ODE course at Monmouth College. It does not take up very much additional class time, since we would go through most of the same steps (e.g. integration by parts) in order to prove the properties of the transform in the standard presentation. The extra time is justified by the following benefits.

- The Laplace transform is seen as one example of an integral transform, rather than as an isolated definition of a clever computational "trick."
- It lays some groundwork for consideration of the inverse transform.
- It ties the ODE course back to linear algebra, and provides an example involving spaces which are not finite-dimensional.

- It reinforces the concept that mathematics is (at least sometimes) a constructive, goal-oriented process.
- There are several opportunities to raise questions which lead students into more advanced courses, such as functional analysis.

The most significant benefit may be that this presentation involves students in the "constructive" phase of the topic. Rather than simply handing them a tool and sending them out to use it computationally, we hope to involve them in the construction/discovery of the tool. Although experience with this presentation is limited to a few sections, we have seen students ask questions such as, "What would we get if the interval was a finite interval?" One exceptional student conjectured that we might get a tool for use in boundary value problems, since the integration by parts would result in an expression involving the solution at both endpoints. Such curiosity (and insight) seems less likely with the definition-first approach common in most texts.