

Operator Calculus

Vijaysekhar Chellaboina

September 17, 2014

Abstract

Today, I was asked an interesting question in my tutorial about mathematical background of the *operator calculus*, the method that involves solving an algebraic equation of *operators* introduced in EE 101 to help solving linear ordinary differential equations. In this note, I would like to connect operator calculus to some of the rigorous techniques we are presenting in MA 101. This note is not trying to claim that there is something wrong with operator calculus but only trying to present the mathematical meaning of the same. The first two sections will present clarification on the notation and the third section will present connections to solutions to linear time-invariant differential equations.

1 Operators: Basics

As we have seen in MA 101, a function, say f , always has a domain D and a range R and we use the notation $f : D \rightarrow R$ to denote that the function takes elements from D to produce elements of R . In MA 101, we assumed D and R are subsets of the real numbers \mathbb{R} .

Now, let \mathcal{F} denote the set of all *smooth* real functions (operating on real numbers). Here, smooth indicates that the function is differentiable everywhere and upto any order (that is, twice differentiable, thrice differentiable etc.). So if $f \in \mathcal{F}$ then $f : \mathbb{R} \rightarrow \mathbb{R}$ is a smooth function.

With a little bit of imagination we can define functions $F : \mathcal{F} \rightarrow \mathcal{F}$. For example, indefinite integrals may be viewed as such functions. Alternatively, definite integrals may be viewed as functions with domain \mathcal{F} and range \mathbb{R} . It is customary to call functions with domain \mathcal{F} are called as *operators*. One such operator is the first derivative. Specifically, let $D : \mathcal{F} \rightarrow \mathcal{F}$ denote an operator defined by $Dy = y'$ for all $y \in \mathcal{F}$. Note that since $y \in \mathcal{F}$ it follows that $y' \in \mathcal{F}$ (since y is smooth). Hence, the notation $D : \mathcal{F} \rightarrow \mathcal{F}$ is well defined. Now, we can define D^2 to denote the composite operator $D \circ D$. It is easy to show that $D^2y = y''$ for all $y \in \mathcal{F}$ and $D^2 : \mathcal{F} \rightarrow \mathcal{F}$. Similarly, we can extend the notation to D^n where n is some nonnegative integer. Specifically, D^0 denotes the identity operator, that is, $D^0y = y$ and D^1 is D . It is customary to denote D^0 by 1.

Now, let L denote a polynomial in D , that is,

$$L = a_0D^0 + a_1D^1 + \cdots + a_nD^n,$$

where $a_k \in \mathbb{R}$, $k = 1, \dots, n$. Once again, it is easy to show that $L : \mathcal{F} \rightarrow \mathcal{F}$, that is, for every $y \in \mathcal{F}$, Ly produces a smooth function. Knowing all the facts of derivatives (introduced in MA 101) it is easy to show that for every pair $y_1, y_2 \in \mathcal{F}$ and $\alpha_1, \alpha_2 \in \mathbb{R}$,

$$L(\alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 Ly_1 + \alpha_2 Ly_2.$$

Hence, L is a *linear operator*. Although this note is about only linear operators, one can imagine nonlinear operators and they pose even more difficult problems.

2 Polynomial Functions of Operators

Ready for next level of abstraction? Let P denote a polynomial function on the set of operators, that is, P is an operator on operators. For example, define $P(D) \equiv D^2 + 2D + 1$. Note that P is not a function on \mathcal{F} or \mathbb{R} or something as simple. It is a function on set of operators and to avoid introducing more notation I chose not give a symbol to this set. P is like a super-operator or a meta-operator. Of course, when P operates on D it produces the *value* $P(D)$ which can be defined as L , an operator from $\mathcal{F} \rightarrow \mathcal{F}$. Though $L = P(D)$ there is a difference between L and $P(D)$. The difference is subtle. For example, $L = D^2 + 2D + 1$ is like saying $b = a^2 + 2a + 1$ for some $a, b \in \mathbb{R}$. But the notation $P(D)$ can be extended to denote some thing like $P(D^2)$, which is $D^4 + 2D^2 + 1$ (more like defining a function $p : \mathbb{R} \rightarrow \mathbb{R}$ given by $p(a) = a^2 + 2a + 1$, which has the flexibility of denoting other values such as $p(a^2)$ etc.). With a minor abuse of notation, I would like to use the notation $P(\lambda)$ (where $\lambda \in \mathbb{R}$) to denote $\lambda^2 + 2\lambda + 1$ (in this example). Why is this an abuse of notation? Because the notation $P(\lambda)$ is valid only if $P : \mathbb{R} \rightarrow \mathbb{R}$ and we defined P as an operator on operators.

Now, we will end this section with an interesting observation. let $y \in \mathcal{F}$ be given by $y(x) = e^{\lambda x}$ for all $x \in \mathbb{R}$ where $\lambda \in \mathbb{R}$. Continuing with the same example, $L = P(D) = D^2 + 2D + 1$,

$$Ly(x) = P(D)(y(x)) = (D^2 + 2D + 1)(e^{\lambda x}) = (\lambda^2 + 2\lambda + 1)e^{\lambda x} = P(\lambda)y(x).$$

Hidden above is the identity $P(D)y(x) = P(\lambda)y(x)$, which will be very useful below.

3 Linear Differential Equations

Let L a polynomial of the operator D , that is, $L = P(D)$ where P is the corresponding polynomial function of operators. Then for a given function $y \in \mathcal{F}$,

$$Ly = P(D)y = 0$$

denotes a *linear time-invariant (ordinary) differential equation*. For example, if $L = D^2 + 2D + 1$ then $Ly = 0$ is same as

$$(D^2 + 2D + D^0)y = 0,$$

or, equivalently,

$$D^2y + 2Dy + y = 0.$$

In the usual notation of derivatives $Ly = 0$ may also be represented by

$$\frac{d^2y}{dx^2} + 2\frac{dy}{dx} + y = 0,$$

where we implicitly assumed that function $y \in \mathcal{F}$ is written in terms of the *independent (real) variable* x , that is, $y(x)$ denotes the value of y at x , $Dy(x)$ denotes $y'(x)$, the first derivative at x etc.

Now, consider the problem of determining solutions to the differential equation

$$Ly = P(D)y = 0,$$

where L is a polynomial of the operator D and P , the corresponding polynomial function. Of course, many issues will pop up immediately: *i*) What do we mean by a solution to a differential equation?, *ii*) Is it right to say “the solution”, that is, can we claim that the solution is unique?. All these questions will be addressed in a later semester.

Since we have shown that $L(e^{\lambda_k x}) = P(\lambda_k)e^{\lambda_k x}$ it follows that $P(\lambda_k) = 0$ for all $k = 1, \dots, n$. So a *standard technique* to solve differential equations of the form $P(D)y = 0$ is to solve (for λ) the algebraic equation

$$P(\lambda) = 0,$$

and use the roots to finally write out the solution to our original differential equation $Ly = 0$. The equation $P(\lambda) = 0$ is called as the *characteristic equation* of the differential equation $Ly = 0$.

Sometimes we are allowed to be *lazy* and treat $P(D) = 0$ as an algebraic equation and roots of the equation can then be used to get the solutions of the differential equation. Of course, mathematically, this is wrong. But if you understand the background then you are allowed (in non-MA courses) to break the rules for sake of *convenience*.

We will close this note by stating a theorem without giving a proof. The proof awaits you in a more advanced mathematics course.

Theorem 1 *Let P denote an n -th order polynomial function of operators. Assume that $P(\lambda) = 0$ have n distinct roots denoted by λ_k , $k = 1, 2, \dots, n$. Then the general solution to the linear time-invariant differential equation $Ly = 0$ (where $L = P(D)$) is given by $\sum_{k=1}^n c_k e^{\lambda_k x}$.*

The case of non-distinct roots is more involved and I am sure it is discussed clearly in EE101 (at least for second order differential equations).

Hope this helps.