Notes on Convolution

Situation: We need to solve a differential equation of the form $p(D)[x(t)] = f(t)$ with initial conditions $x(0) = x_0$, $\dot{x}(0) = \dot{x}_0$, etc.

Plan: If we can find the unit impulse response for this system (with rest initial conditions), i.e. the solution to $p(D)[x(t)] = \delta(t)$ with initial conditions $x(0) = 0,\ \dot{x}(0) = 0$, etc., we will develop a method for finding a solution to $p(D)[x(t)] = f(t)$ by thinking of $f(t)$ as a "train of impulses." We will likely use Laplace transform methods to find the unit impulse response. We'll use Riemann Sums ideas to construct an integral by piecing together solutions associated with the impulses. This will be the **convolution** integral.

Managing the Initial Conditions

Given a "driven" system governed by an ODE such as $p(D)[x(t)] = f(t)$ with initial conditions $x(t_0) = x_0$, $\dot{x}(t_0) = \dot{x}_0$, etc., we generally identify the left-hand expression $p(D)[x(t)]$ as "the system" and the inhomogeneity on the right-hand side $f(t)$ as the "input signal". There is some useful terminology relevant to these types of ODE's.

If $x(t_0) = 0$, $\dot{x}(t_0) = 0$, etc., we refer to this as the **zero state**.

Solving $p(D)[x(t)] = f(t)$ for the zero state, we refer to the solution $x_f(t)$ as the **zero state response (ZSR)**.

If we seek homogeneous solutions to the ODE $p(D)[x(t)] = 0$ for any state with $x(t_0) = x_0$ and $\dot{x}(t_0) = \dot{x}_0$, this will have a unique solution $x_h(t)$ called the **zero input response (ZIR)**.

In general, the solution to the ODE $p(D)[x(t)] = f(t)$ will be $x(t) = x_h(t) + x_p(t)$ for some particular solution $x_p(t)$, but note that the zero state response (**ZSR**) is such a particular solution, so $x(t) = x_h(t) + x_f(t)$. That is, $x(t) = \textbf{ZIR} + \textbf{ZSR}$. Note that $\left\{\begin{matrix} x(t_0) = x_h(t_0) + x_f(t_0) = x_h(t_0) + 0 = x_h(t_0) = x_0 \\ \dot{x}(t_0) = \dot{x}_h(t_0) + \dot{x}_f(t_0) = \dot{x}_h(t_0) + 0 = \dot{x}_h(t_0) = \dot{x}_0 \\ \text{etc.} \end{matrix}\right\}$, so $x(t) = x_h(t) + x_p(t)$ satisfies the initial value problem (IVP) without the need to introduce any additional constants.

Developing the convolution integral

(1) We start by solving for the unit impulse function $w(t)$, i.e. the solution to $p(D)[x(t)] = \delta(t)$ with initial conditions $x(0) = 0,\ \dot{x}(0) = 0$, etc. We refer to this by $w(t)$ and reserve $x(t)$ for the solution to $p(D)[x(t)] = f(t)$. If $W(s)$ is the Laplace transform of $w(t)$, we'll have $p(s)W(s) = 1$ where $p(s)$ is the characteristic polynomial, so $W(s) = \dfrac{1}{p(s)}$ and with some partial fractions calculations and the inverse Laplace transform, finding $w(t)$ can be reduced to a relatively simple routine.
We call $w(t)$ the ***weight function***, and we call $W(s)$ the ***transfer function***.

(2) We use **time invariance** to declare that the translated unit impulse response for $p(D)[x(t)] = \delta(t - t_k)$ will be $w(t - t_k)$. We'll use this in the integral to follow.

(3) If we're interested in understanding what's happening during the time interval $[0, t]$, we start by partitioning this interval into many small subintervals $[t_{k-1}, t_k]$. On each of these subintervals, we can use a box function to "switch on" just one small section of the function $f(t)$. That is, if we let $f_k(t) = f(t)[u(t - t_{k-1}) - u(t - t_k)]$ then this function will be identically zero except in the k-th subinterval $[t_{k-1}, t_k]$. We can later reassemble the function as $f(t) = \sum f_k(t)$.

If $f(t)$ is reasonably well behaved (except, perhaps, at finitely many points), we can say that within a given subinterval, $f_k(t) \cong f(t_k) = \int_0^{+\infty} f(t)\delta(t-t_k)dt = \int_{-\infty}^{+\infty} f(t)\delta(t-t_k)dt$, where we use the fact that evaluation of a function at a point is accomplished by integrating against a delta function concentrated at that point.

(4) If the solution to $p(D)[x(t)] = \delta(t-t_k)$ is $w(t-t_k)$, then linearity gives that the solution to $p(D)[x(t)] = f(t_k)\delta(t-t_k)\Delta t_k$ will be $f(t_k)w(t-t_k)\Delta t_k$ where $\Delta t_k = t_k - t_{k-1}$ is the width of the k-th subinterval. [We simply multiplied both sides by $f(t_k)\Delta t_k$.]

(5) By linearity (superposition), we can sum to get that the solution to $p(D)[x(t)] \cong \sum_k f(t_k)\delta(t-t_k)\Delta t_k$ must therefore be $x(t) \cong \sum_k f(t_k)w(t-t_k)\Delta t_k = \sum_k f(t_k)w(t-\tau_k)\Delta\tau_k$, where we changed to the variable τ in anticipation of the next step.

(6) If we pass to the limit as the norm of the partition goes to zero, the sum will become an integral and the approximation will become exact, i.e. $x(t) = \int_0^t f(\tau)w(t-\tau)d\tau \equiv (f*w)(t)$, the convolution integral. This provides a solution to $p(D)[x(t)] = f(t)$ for the zero state (rest conditions), and we refer to this solution as the **zero state response (ZSR)**.

Definition (Convolution): Given two functions $w(t)$ and $f(t)$, we define $\boxed{(f*w)(t) = \int_0^t f(\tau)w(t-\tau)d\tau}$.

It's a straightforward exercise to show that the convolution product is commutative, i.e. $f*w = w*f$.

The fact that the solution to the differential equation $p(D)x = f(t)$ will have the solution $(f*w)(t)$ is also known as ***Green's Formula***.

Note: When applying the convolution method to solving $p(D)x = f(t)$ for more general initial conditions, the solution will be $x(t) = \mathbf{ZIR} + \mathbf{ZSR}$, where **ZIR** is the **zero input response** and **ZSR** is the **zero state response**.

The usefulness of this transform method is built on the fact that we can relatively easily find the Laplace transform for most everything that appears in a given differential equation of the form $p(D)x = f(t)$, and once we have a table of these transforms we can generally invert the process by inspection. Another essential fact is that the Laplace transform acts linearly, and this allows us to decompose complex problems into a sums of simple problems.

Example (with a familiar input): Solve $\ddot{x} + 3\dot{x} + 2x = 2e^{-t}, \quad x(0) = 0, \quad \dot{x}(0) = 0$ using (a) previous methods, (b) using only the Laplace transform, and (c) using the Laplace transform plus convolution.

Solution: **(a)** For the homogeneous equation $\ddot{x} + 3\dot{x} + 2x = 0$, the characteristic polynomial is $p(s) = s^2 + 3s + 2 = (s+2)(s+1)$ which yields the two roots $s = -2$ and $s = -1$. This gives the two independent solutions e^{-2t} and e^{-t}, and all homogeneous solutions are of the form $x_h(t) = c_1e^{-2t} + c_2e^{-t}$.

In seeking a particular solution $x_p(t)$ that satisfies the inhomogeneous differential equation, we see that the Exponential Response Formula (ERF) won't work – there is resonance. We can, however, use the Resonant Response Formula to get the particular solution $x_p(t) = \frac{2te^{-t}}{p'(-1)} = \frac{2te^{-t}}{1} = 2te^{-t}$, so the general solution is $x(t) = x_h(t) + x_p(t) = c_1e^{-2t} + c_2e^{-t} + 2te^{-t}$. Its derivative is $\dot{x}(t) = -2c_1e^{-2t} - c_2e^{-t} - 2te^{-t} + 2e^{-t}$. Substituting the (rest) initial conditions gives $\left\{\begin{matrix} x(0) = c_1 + c_2 = 0 \\ \dot{x}(0) = -2c_1 - c_2 + 2 = 0 \end{matrix}\right\}$, and these can be solved to give $c_1 = 2, c_2 = -2$, so the solution is $\boxed{x(t) = 2e^{-2t} - 2e^{-t} + 2te^{-t}}$.

(b) We need the following Laplace transforms:

(1) $\mathcal{L}(e^{kt}) = \dfrac{1}{s-k}$ with region of convergence $\mathrm{Re}(s) > k$, so $\mathcal{L}(e^{-2t}) = \dfrac{1}{s+2}$.

(2) If the Laplace transform of $x(t)$ is $X(s)$, then the Laplace transforms of its derivatives are $\mathcal{L}(\dot{x}(t)) = sX(s) - x(0-)$ and $\mathcal{L}(\ddot{x}(t)) = s^2X(s) - s\,x(0-) - \dot{x}(0-)$. We have rest initial conditions, so these are greatly simplified and, in fact, $\mathcal{L}(p(D)x) = p(s)X(s)$. Specifically, $\mathcal{L}(\ddot{x} + 3\dot{x} + 2x) = s^2X(s) + 3sX(s) + 2X(s) = (s^2 + 3s + 2)X(s) = p(s)X(s)$.

If we now transform the entire differential equation, we get $(s^2 + 3s + 2)X(s) = \dfrac{2}{s+1}$.

We then solve for $X(s) = \dfrac{2}{(s+1)(s^2+3s+2)} = \dfrac{2}{(s+2)(s+1)^2} = \dfrac{A}{s+2} + \dfrac{B}{s+1} + \dfrac{C}{(s+1)^2}$.

There are many good ways to find the unknowns *A*, *B*, and *C*. For example, if we multiply through by the common denominator to clear fractions, we get $2 = A(s+1)^2 + B(s+1)(s+2) + C(s+2)$. Plugging in the specific values $s = -2$ and $s = -1$ quickly yields that $A = 2$ and $C = 2$. Plugging in, for example, $s = 0$ and using the values for *A* and *C* then yields $B = -2$. So $X(s) = \dfrac{2}{s+2} - \dfrac{2}{s+1} + \dfrac{2}{(s+1)^2}$.

Consulting our table of common Laplace transforms, we see that $\dfrac{2}{s+2} = \mathcal{L}(2e^{-2t})$, $\dfrac{2}{s+1} = \mathcal{L}(2e^{-t})$, and $\dfrac{2}{(s+1)^2} = \mathcal{L}(2te^{-t})$, so transforming back gives $\boxed{x(t) = 2e^{-2t} - 2e^{-t} + 2te^{-t}}$.

(c) We start by finding the unit impulse response, a the solution to $\ddot{x} + 3\dot{x} + 2x = \delta(t)$ with rest initial conditions $x(0) = 0, \dot{x}(0) = 0$. Laplace transform gives $p(s)W(s) = 1$, so $W(s) = \dfrac{1}{(s+2)(s+1)} = -\dfrac{1}{s+2} + \dfrac{1}{s+1}$. Consulting the Laplace transform table, this yields the weight function $w(t) = \left\{\begin{matrix} 0 & t < 0 \\ -e^{-2t} + e^{-t} & t > 0 \end{matrix}\right\} = u(t)(-e^{-2t} + e^{-t})$.

With $f(t) = 2e^{-t}$, then convolution of the weight function and the given input signal gives:

$$\begin{aligned}(w * f)(t) &= \int_{\tau=0}^{\tau=t} w(\tau) f(t-\tau)\,d\tau = \int_{\tau=0}^{\tau=t} (-e^{-2\tau} + e^{-\tau})2e^{-(t-\tau)}\,d\tau \\ &= \int_{\tau=0}^{\tau=t} 2e^{-t}(-e^{-\tau} + 1)\,d\tau = 2e^{-t}\left[e^{-\tau} + \tau\right]_{\tau=0}^{\tau=t} = 2e^{-t}[e^{-t} - 1 + t] \\ &= \boxed{+2e^{-2t} - 2e^{-t} + 2te^{-t} = x(t)}\end{aligned}$$

Alternatively, we could have calculated this as:

$$\begin{aligned}(f * w)(t) &= \int_{\tau=0}^{\tau=t} f(\tau) w(t-\tau)\,d\tau = \int_{\tau=0}^{\tau=t} 2e^{-\tau}(-e^{-2(t-\tau)} + e^{-(t-\tau)})\,d\tau \\ &= \int_{\tau=0}^{\tau=t} (-2e^{-2t}e^{\tau} + 2e^{-t})\,d\tau = -2e^{-2t}(e^{t} - 1) + 2e^{-t}(t-0) = \boxed{+2e^{-2t} - 2e^{-t} + 2te^{-t} = x(t)}\end{aligned}$$

Example: Solve the same ODE as above but with non-rest initial conditions:

$$\ddot{x}+3\dot{x}+2x=2e^{-t}, \quad x(0)=4, \quad \dot{x}(0)=0$$

Solution: All of the previous steps are the same in deriving the **zero state response (ZSR)**, so we have

$$\textbf{ZSR}=+2e^{-2t}-2e^{-t}+2te^{-t}.$$

We need only find the **zero input response (ZIR)**. This simply means that we solve $\ddot{x}+3\dot{x}+2x=0$ to get

$$x_h(t)=c_1e^{-2t}+c_2e^{-t} \text{ and } \dot{x}_h(t)=-2c_1e^{-2t}-c_2e^{-t}, \text{ so } \left\{\begin{array}{l} x_h(0)=c_1+c_2=4 \\ \dot{x}_h(0)=-2c_1-c_2=0 \end{array}\right\} \Rightarrow c_1=-4,\ c_2=8.$$

So $\textbf{ZIR}=-4e^{-2t}+8e^{-t}$.

Therefore $x(t)=\textbf{ZSR}+\textbf{ZIR}=+2e^{-2t}-2e^{-t}+2te^{-t}-4e^{-2t}+8e^{-t}=\boxed{-2e^{-2t}+6e^{-t}+2te^{-t}}$.