1. First-Order Linear Ordinary Differential Equations with Constant Coefficients

In this course, we will learn how to solve the differential equations and let's start with the simplest case, which is the 1st-order linear ordinary differential equation with constant coefficients. In general, it is expressed as

(1-1) $$L[y(t)] \equiv p_1\dot{y}(t) + p_0 y(t) = Q(t)$$

where $L[y(t)] \equiv p_1\dot{y}(t) + p_0 y(t)$, *t* is an independent variable within an interval *I*, $Q(t)$ is a given function and $y(t)$ is the unknown function to be solved. Due to the fact that differential equations are often used to analyze practical systems, we will focus on the case that p_0, p_1, *t*, $Q(t)$ and $y(t)$ are all real numbers.

In mathematics, a differential equation with functions of one independent variable is called an ordinary differential equation (ODE), while a differential equation involving functions of more than one independent variable is called a partial differential equation (PDE). From now on, we shall pay attention to the ODEs.

Because only one independent variable *t* appears in (1-1) and the highest-order derivative is the first derivative $\dot{y}(t)$, we say that (1-1) a 1st-order ODE. Moreover, since $L[y(t)] \equiv p_1\dot{y}(t) + p_0 y(t)$ is a linear combination of $y(t)$ and $\dot{y}(t)$ with constant coefficients p_1 and p_0, (1-1) is further called as a 1st-order linear ODE with constant coefficients. For the sake of brevity again, the term "ODE with constant coefficients" will be shortly denoted as "CODE". In summary, (1-1) is said to be a 1st-order linear CODE.

1st-order Linear CODEs

Without loss of generality, the coefficient $p_1 \neq 0$ in (2) is often normalized to 1. Such normalization process is called "monic". Through the monic process, (1-1) can be changed into the following normal form

(1-2) $$\dot{y}(t) + p y(t) = q(t)$$

where $p = \frac{p_0}{p_1}$ and $q(t) = \frac{Q(t)}{p_1}$. Next, let's discuss two properties related to the solution $y(t)$ in (1-2), called existence and uniqueness.

For the existence of $y(t)$, instead of mathematical proof, we will just find out a solution to declare its existence. Commonly, a solution of (1-2) can be solved by setting

(1-3) $$y(t) = z(t)e^{-pt}$$

whose derivative is

(1-4) $$\dot{y}(t) = \dot{z}(t)e^{-pt} - py(t)$$

In comparison with (1-2), we know that

(1-5) $$\dot{y}(t) + py(t) = \dot{z}(t)e^{-pt} = q(t)$$

which results in

(1-6) $$\dot{z}(t) = q(t)e^{pt}$$

Clearly, taking integration will obtain

(1-7) $$z(t) = \int_a^t q(\tau)e^{p\tau}d\tau + z(a)$$

where a is an arbitrary constant. From (1-3), there indeed exists a solution, usually called a particular solution and denoted by $y_p(t)$, which is shown as

(1-8)
$$\begin{aligned} y_p(t) &= z(t)e^{-pt} = e^{-pt}\int_a^t q(\tau)e^{p\tau}d\tau + z(a)e^{-pt} \\ &= \int_a^t q(\tau)e^{-p(t-\tau)}d\tau + z(a)e^{-pa}e^{-p(t-a)} \\ &= \int_a^t q(\tau)e^{-p(t-\tau)}d\tau + y_p(a)e^{-p(t-a)} \end{aligned}$$

with $y_p(a) = z(a)e^{-pa}$. Hence, the existence of the solution of (1-2) is guaranteed.

Homogeneous Equation

As for the uniqueness of $y(t)$ in (1-2), it is easy to check that $y_p(t)$ in (1-8) is not unique since a is an arbitrary constant. In other words, a should be fixed if $y_p(t)$ is unique. Now, one question is raised: On what condition $y_p(t)$ is unique? Before answering the question, let's discuss the case of $q(t)=0$ in (1-2), i.e.,

(1-9) $$\dot{y}(t) + py(t) = 0$$

which is known as a homogeneous equation.

There are several methods able to solve a 1st-order homogeneous equation. For example, the method by the use of (1-3), which has been introduced for the case of $q(t)\neq 0$, is also available for the homogeneous equation (1-9). However, these methods may not be usable for higher-order homogeneous equations.

Here, we will focus on a method, which is not only suitable for 1st-order homogeneous equations, but also extendable to homogeneous equations of higher order. First, let's denote the solution of a homogeneous equation as $y_h(t)$ and call it the homogeneous solution. Then, we assume

(1-10) $$y_h(t) = e^{\lambda t}$$

where λ is a constant. Further substitute it into (1-9) and achieve $\lambda e^{\lambda t} + pe^{\lambda t} = 0$. Since $e^{\lambda t} \neq 0$, we obtain

(1-11) $$\lambda + p = 0$$

which is called the characteristic equation of (1-9). Clearly, the root, or formally the characteristic root, is $\lambda = -p$. From (1-10), the homogeneous solution should be in the form of $y_h(t) = e^{-pt}$ and thus we can choose a more general form as

(1-12) $$y_h(t) = A_h e^{-pt}$$

where A_h is an arbitrary constant.

Based on the above analysis, the solutions $y_p(t)$ and $y_h(t)$ satisfy the following equations

(1-13) $$\dot{y}_p(t) + py_p(t) = q(t)$$

(1-14) $$\dot{y}_h(t) + py_h(t) = 0$$

Combining them together, we have

(1-15) $$\frac{d}{dt}\left(y_p(t) + y_h(t)\right) + p\left(y_p(t) + y_h(t)\right) = q(t)$$

Obviously, $y(t) = y_p(t) + y_h(t)$ is also a solution of (1-2) and expressed as

(1-16)
$$\begin{aligned} y(t) &= y_p(t) + y_h(t) \\ &= \int_a^t q(\tau)e^{-p(t-\tau)}d\tau + y_p(a)e^{-p(t-a)} + A_h e^{-pt} \\ &= \int_a^t q(\tau)e^{-p(t-\tau)}d\tau + Ae^{-p(t-a)} \\ &= \int_a^t q(\tau)e^{-p(t-\tau)}d\tau + y(a)e^{-p(t-a)} \end{aligned}$$

where $A = y(a) = y_p(a) + A_h e^{-pa}$ is also an arbitrary constant.

To sum up, the general solution $y(t) = y_p(t) + y_h(t)$ of a 1st-order linear CODE is composed of a particular solution $y_p(t)$ and a homogeneous solution $y_h(t)$. However, it is still not unique since A is an arbitrary constant. In order to get a unique solution, it is required to add an extra condition, such as an initial condition or a boundary condition. Next, let's discuss the initial value problem, which includes an initial condition.

Initial Value Problem

To get a unique solution for (1-16), an extra condition is required to determine the value of A. For example, consider an interval $I=[t_0,\infty)$ and $t\in I$, i.e., $t \ge t_0$ and t_0 is the initial point of the interval. Let $y(t_0)=y_0$ be an extra condition, which is called the initial condition. Then, (1-2) can be rewritten as

(1-17) $$\dot{y}(t)+py(t)=q(t), \quad y(t_0)=y_0$$

which is known as an initial value problem, or IVP in short. Now, we can assign $a=t_0$ for (1-16) and obtain

(1-18) $$y(t)=\int_{t_0}^{t} q(\tau)e^{-p(t-\tau)}d\tau + y_0 e^{-p(t-t_0)}, \quad \text{for } t\ge t_0$$

Obviously, (1-18) is the unique solution of the IVP shown in (1-17).

In engineering, IVPs are the most common problems to be dealt with. Figure 1-1 shows an example of IVPs, which is an RC circuit designed to reduce the higher-frequency signals generated by the input voltage source $v_s(t)$, such that only the signals in $v_s(t)$ of lower-frequency can pass through the circuit. That means the output capacitor voltage $v_C(t)$ will only contains the lower-frequency signals generated by $v_s(t)$. This circuit is a kind of 1st-order low-pass filter.

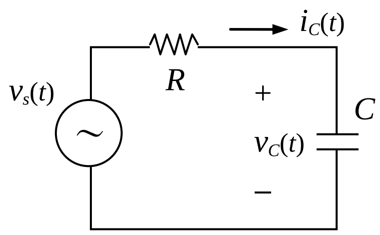


Figure 1-1

In the RC circuit, the capacitor is initially possessed of a voltage $v_C(0)=v_{C0}$ at the initial time t=0. Based the Kirchhoff's voltage law, the dynamic model of the circuit is obtained as

(1-19) $$v_s(t)=Ri_C(t)+v_C(t)=RC\dot{v}_C(t)+v_C(t)$$

where the capacitor current is given as $i_C(t)=C\dot{v}_C(t)$. Rearrange (1-19) as

(1-20) $$\dot{v}_C(t)+pv_C(t)=q(t), \quad v_C(0)=v_{C0}$$

where $p=\frac{1}{RC}$ and $q(t)=\frac{v_s(t)}{RC}$. Clearly, this is an IVP and according to (1-18) the solution is determined as

(1-21) $$v_C(t)=\frac{1}{RC}\int_0^t v_s(\tau)e^{-(t-\tau)/RC}d\tau+v_{C0}e^{-t/RC}, \quad \text{for } t\ge 0$$

which is the unique solution of the IVP given in (1-20).

Particular Solutions

Now, let's consider the IVP of a 1st-order linear CODE shown in (1-17) with some given functions $q(t)\in\{1,t,t^2,sin\,\omega t,cos\,\omega t\}$. Let the initial point be t=0, then the IVP is expressed as

(1-22) $$\dot{y}(t)+py(t)=q(t), \qquad y(0)=y_0$$

where $y(0)=y_0$ is the initial condition, and the unique solution is

(1-23) $$y(t)=\int_0^t q(\tau)e^{-p(t-\tau)}d\tau+y_0e^{-pt}$$

which is composed of a particular solution given in (1-8) as

(1-24) $$y_p(t)=\int_0^t q(\tau)e^{-p(t-\tau)}d\tau+y_p(0)e^{-pt}$$

and a homogeneous solution obtained by

(1-25) $$y_h(t)=y(t)-y_p(t)=(y_0-y_p(0))e^{-pt}$$

Next, let's focus on the particular solution corresponding to each of the given functions $q(t)\in\{1,t,t^2,sin\,\omega t,cos\,\omega t\}$.

The first function is $q(t)=1$. From (1-24), the particular solution is calculated as below:

(1-26) $$\begin{aligned} y_p(t)&=e^{-pt}\left(\int_0^t e^{p\tau}d\tau+y_p(0)\right)=e^{-pt}\left(\frac{1}{p}e^{p\tau}\Big|_{\tau=0}^{t}+y_p(0)\right)\\ &=\frac{1}{p}\left(1-e^{-pt}\right)+y_p(0)e^{-pt}=\frac{1}{p}+\left(y_p(0)-\frac{1}{p}\right)e^{-pt}\end{aligned}$$

Because only one particular solution is required, we choose $y_p(0)=\frac{1}{p}$ and get the simplest one $y_p(t)=\frac{1}{p}$, which is constant for all t. In other words, if $q(t)$ is constant, the particular solution $y_p(t)$ can be also chosen as a constant.

For the case of $q(t)=t$, the particular solution can be also calculated from (1-24) as below:

(1-27)
$$y_p(t)=e^{-pt}\left(\int_0^t \tau e^{p\tau}d\tau+y_p(0)\right)=e^{-pt}\frac{1}{p}\left(\tau e^{p\tau}\Big|_{\tau=0}^{t}-\int_0^t e^{p\tau}d\tau+y_p(0)\right)$$
$$=e^{-pt}\left(\frac{1}{p}te^{pt}-\frac{1}{p^2}\left(e^{pt}-1\right)+y_p(0)\right)=\frac{1}{p}t-\frac{1}{p^2}+\left(\frac{1}{p^2}+y_p(0)\right)e^{-pt}$$

We choose the simplest one $y_p(t)=\frac{1}{p}t-\frac{1}{p^2}$. It is clear that the term e^{-pt} vanishes since $y_p(0)=-\frac{1}{p^2}$. Hence, if $q(t)$ is a 1st-order polynomial, then the particular solution $y_p(t)$ can be also chosen as a 1st-order polynomial.

Similarly, if $q(t)=t^2$ is a 2nd-order polynomial, then the particular solution can be chosen as

(1-28)
$$y_p(t)=\frac{1}{p}t^2-\frac{2}{p^2}t+\frac{2}{p^3}$$

which is also a 2nd-order polynomial.

To find the particular solutions for $q(t)=sin\,\omega t$ and $q(t)=cos\,\omega t$, let's use the case of $q(t)=e^{j\omega t}=cos\,\omega t+j\,sin\,\omega t$ as a substitution. From (1-24), we obtain

(1-29)
$$y_p(t)=e^{-pt}\left(\int_0^t e^{j\omega\tau}e^{p\tau}d\tau+y_p(0)\right)=e^{-pt}\left(\frac{1}{p+j\omega}e^{(p+j\omega)\tau}\Big|_{\tau=0}^{t}+y_p(0)\right)$$
$$=\frac{1}{p+j\omega}e^{j\omega t}+\left(y_p(0)-\frac{1}{p+j\omega}\right)e^{-pt}$$

The particular solution is chosen as $y_p(t)=\frac{1}{p+j\omega}e^{j\omega t}$, which makes the term e^{-pt} vanish since $y_p(0)=\frac{1}{p+j\omega}$. Notice that the particular solution of $q(t)=e^{j\omega t}$ can be further written into

(1-30)
$$y_p(t)=\frac{1}{p+j\omega}e^{j\omega t}=\frac{p-j\omega}{p^2+\omega^2}\left(cos\,\omega t+j\,sin\,\omega t\right)$$
$$=\frac{1}{p^2+\omega^2}\left(p\,cos\,\omega t+\omega\,sin\,\omega t\right)+j\frac{1}{p^2+\omega^2}\left(p\,sin\,\omega t-\omega\,cos\,\omega t\right)$$

Correspondingly, if $q(t)=sin\,\omega t$, which is the imaginary part of $e^{j\omega t}$, then its particular solution is the imaginary part of (1-30), expressed as

(1-31) $$y_p(t)=\frac{1}{p^2+\omega^2}\left(p\sin\omega t-\omega\cos\omega t\right)$$

Similarly, if $q(t)=\cos\omega t$, which is the real part of $e^{j\omega t}$, then its particular solution is the real part of (1-30), expressed as

(1-32) $$y_p(t)=\frac{1}{p^2+\omega^2}\left(p\cos\omega t+\omega\sin\omega t\right)$$

The solutions of the IVP given in (1-22) with $q(t)\in\{1,t,t^2,\sin\omega t,\cos\omega t\}$ are listed in Table 1-1.

Table 1-1

$q(t)$	Solution of $\dot{y}(t)+py(t)=q(t),\ \ y(0)=y_0$
1	$y(t)=\left(y_0-\frac{1}{p}\right)e^{-pt}+\frac{1}{p}$
t	$y(t)=\left(y_0+\frac{1}{p^2}\right)e^{-pt}+\frac{1}{p}t-\frac{1}{p^2}$
t^2	$y(t)=\left(y_0-\frac{2}{p^3}\right)e^{-pt}+\frac{1}{p}t^2-\frac{2}{p^2}t+\frac{2}{p^3}$
$\sin\omega t$	$y(t)=\left(y_0+\frac{\omega}{p^2+\omega^2}\right)e^{-pt}+\frac{1}{p^2+\omega^2}\left(p\sin\omega t-\omega\cos\omega t\right)$
$\cos\omega t$	$y(t)=\left(y_0-\frac{p}{p^2+\omega^2}\right)e^{-pt}+\frac{1}{p^2+\omega^2}\left(p\cos\omega t+\omega\sin\omega t\right)$

Time Constant

Here, we will discuss an important concept called the time constant of a physical system, whose dynamic equation is described by the following 1st-order linear CODE:

(1-33) $$\dot{y}(t)+py(t)=q(t),\quad y(t_0)=y_0$$

Note that *t* is the variable of time and t_0 is the initial time. Referring to (1-23), (1-24) and (1-25), the solution can be expressed as

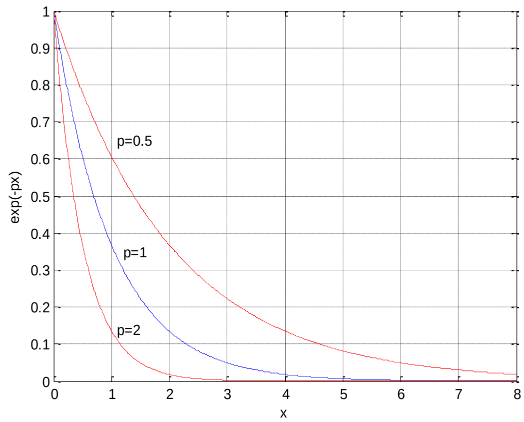
(1-34) $$y(t)=\left(y_0-y_p(t_0)\right)e^{-p(t-t_0)}+y_p(t)$$

where $y_h(t)=\left(y_0-y_p(t_0)\right)e^{-p(t-t_0)}$ is the homogeneous solution and $y_p(t)$ is a particular solution. It is clear that if $p>0$ then

(1-35) $$\begin{cases} y_h(t) = \left(y_0 - y_p(t_0)\right)e^{-p(t-t_0)} \to 0 \\ y(t) = y_h(t) + y_p(t) \to y_p(t) \end{cases}$$

as $t \to \infty$. In general, a physical system satisfying (1-35) is said to be stable. That means the homogeneous solution $y_h(t) = \left(y_0 - y_p(t_0)\right)e^{-p(t-t_0)}$ of a stable system will decrease as *t* increases. The trend of decreasing in time is due to the term e^{-pt} with *p*>0. In what follows, the Matlab program figexp.m is used to plot the function e^{-px} for the cases of *p*=0.5, 1 and 2 during the interval $x \in [0,8]$. The numerical results are shown in the figure. From the three curves corresponding to *p*=0.5, 1 and 2, it is true that $e^{-px} \to 0$ as $x \to \infty$.

```
figexp.m
    clear
    % plot y(x) = exp(-px) for p=0.5,1,2
    x = 0:0.01:8;
    y1 = exp(-x);      % p=1
    y2 = exp(-0.5*x);    % p=0.5
    y3 = exp(-2*x);    % p=2
    plot(x,y1,'b',x,y2,'r',x,y3,'r')
    grid, xlabel('x'), ylabel('exp(-px)')
    text(1.1,0.65,'p=0.5')
    text(1.2,0.35,'p=1')
    text(1.1,0.14,'p=2')
```



For the time constant, denoted as τ, it is defined as the time $t=\tau$ at which the convergence rate is $e^{-p\tau} = e^{-1} = 0.3679$. That means the time constant is obtained by setting $p\tau=1$ or

(1-36) $$\tau = \frac{1}{p}$$

At $t=\tau$, we know that the effect caused by the homogeneous solution $y_h(t) = Ae^{-pt}$ will be reduced by the ratio of $e^{-1} = 0.3679$, as shown below:

(1-37) $$y_h(t+\tau) = Ae^{-p(t+\tau)} = e^{-1}Ae^{-pt} = e^{-1}y_h(t)$$

Now, one question is raised: What is the amount of time required to neglect the effect of $y_h(t) = \left(y_0 - y_p(t_0)\right)e^{-p(t-t_0)}$, or the effect caused by the initial condition y_0? The answer is 4τ. The reason to choose four time constants is because after 4τ the term e^{-pt} will be decreased to $e^{-4} = 0.0183$, which has been less than 2% and considered to be negligible in most of the engineering problems.