Signals and Systems

Stage 3: Electrical, Electronic and Communications Eng.

Dr Paul Curran

Room 145, Engineering.

paul.curran@ucd.ie

+353-1-7161846

Overview of Topics

Mathematical background and problem statement.

Laplace transform and solution of linear constant-coefficient ordinary differential equations.

Special case of steady-state response to sinusoidal input. Fourier transform.

Solution of linear constant-coefficient partial differential equations. Fourier series.

Discrete-time systems.

Section 1 – Laplace Transform

- Laplace Transform
- Solving linear constant coefficient ordinary differential equations

Laplace Transform

1749-1827

Given signal f let

$$F(s) = \int_{0^-}^{\infty} f(t)\exp(-st)dt \quad , \quad s \in C$$

Definition 1.1: Given signal f the function, $F(s)$, defined as above for all s for which the integral converges and by analytic continuation for all other s, is called the *Laplace Transform* of f, denoted $\mathcal{L}(f)$.

Note

Pierre Simon Marquis de Laplace rose from relatively humble beginnings to Marquis in 1817.

The lower limit of the integral refers to the limit as one approaches the point $t = 0$ from below. As we shall see we take this unusual limit to permit us to deal with discontinuities at 0, *impulses* at 0 and initial conditions in a relatively elementary manner. Strictly the resulting transform is called the *single-sided* or *unilateral Laplace transform*.

Example 1.1

Definition 1.2: The *step function* $u(t)$ is given by

$$u(t) = \begin{cases} 0 & \text{if } t < 0 \\ \frac{1}{2} & \text{if } t = 0 \\ 1 & \text{if } t > 0 \end{cases}$$

$$U(s) = \frac{1}{s}$$

Note

The definition of the Laplace transform given above is quite technical courtesy of the term “analytic continuation”. Although the concept of analytic continuation is quite subtle the idea is commonly quite straightforward in practice. The case of the unit step is typical. The integral establishes the form of the Laplace transform for those s for which the integral converges, i.e. for s in the open right half plane, i.e. for s having positive real part. Analytic continuation simply declares this to be the solution for all s.

It should be noted that the majority of textbooks *do not take this approach to the definition of the Laplace Transform*. Instead they define the transform to equal the integral for those values of s for which the integral converges and do not define it for other values of s, a rather poor alternative in my opinion.

Laplace Transform Properties: 1

If $\mathcal{L}(f_1(t)) = F_1(s)$ and $\mathcal{L}(f_2(t)) = F_2(s)$

then

$$\mathcal{L}(\alpha_1 f_1(t) + \alpha_2 f_2(t)) = \alpha_1 F_1(s) + \alpha_2 F_2(s)$$

This property is called *linearity*.

Laplace Transform Properties: 2

If $\mathcal{L}(f(t)) = F(s)$ then

$$\mathcal{L}(f(\alpha t)) = \frac{1}{\alpha} F\left(\frac{s}{\alpha}\right) \quad \text{for any } \alpha > 0.$$

This property is called *scaling*.

Laplace Transform Properties: 3

If $\mathcal{L}(f(t)) = F(s)$ then

$$\mathcal{L}(\exp(-\alpha t) f(t)) = F(s+\alpha)$$

If $\mathcal{L}(f(t)) = F(s)$ then

$$\mathcal{L}(f(t-t_0)u(t-t_0)) = \exp(-st_0)F(s)$$

These are called the *shifting* properties.

Laplace Transform Properties: 4

If $\mathcal{L}(f(t)) = F(s)$ then

$$\mathcal{L}\left(\frac{d^n}{dt^n} f(t)\right) = s^n F(s) - s^{n-1} f(0^-) - s^{n-2} f^{(1)}(0^-) - \cdots - f^{(n-1)}(0^-)$$

$$\mathcal{L}(-t\, f(t)) = \frac{d\,F(s)}{ds}$$

These are called the *differentiation* properties.

Laplace Transform Properties: 5

If $\mathcal{L}\left(f(t)\right)=F(s)$ then

$$\mathcal{L}\left(\int_{0^-}^{t} f(\tau)\,d\tau\right)=\frac{F(s)}{s}$$

$$\mathcal{L}\left(\frac{f(t)}{t}\right)=\int_{s}^{\infty} F(s)\,ds$$

These are called the *integration* properties.

Example: 1.2

We apply these properties by reconsidering the liquid-level system.

$$\dot{\tilde{h}} + \alpha \tilde{h} = \gamma \tilde{q}_{in}(t) \quad , \quad t \geq 0.$$

$$\mathcal{L}\left(\dot{\tilde{h}} + \alpha \tilde{h}\right) = s\tilde{H} - \tilde{h}(0) + \alpha \tilde{H} = \gamma \tilde{Q}_{in}(s).$$

$$\tilde{H} = \frac{\tilde{h}(0)}{s+\alpha} + \frac{\gamma \tilde{Q}_{in}(s)}{s+\alpha}.$$

The Laplace transform has transformed the problem from an ordinary differential to an algebraic equation which can be solved (admittedly for the wrong thing).

Note

In order to have a procedure for solving ordinary differential equations we will have to rectify this situation. We will need to have a formula for going backwards and, given the Laplace transform of a signal producing as a result the actual signal itself. In effect this procedure will achieve what Lerch's cancellation law achieved in the introduction.

Inverse Laplace Transform

Definition 1.3: Given $F(s)$ the function $f(t)$ satisfying $f(t) = 0$ for $t < 0$ and $\mathcal{L}(f(t)) = F(s)$ is called the inverse Laplace transform of F, denoted $\mathcal{L}^{-1}(F)$.

Note

If $f_1(t) = f_2(t)$ for $t \geq 0$ then $F_1(s) = F_2(s)$ regardless of how the signals behave for $t < 0$. Hence a function is not uniquely defined over all time by its Laplace transform, since it can behave any way it chooses for negative time. It transpires, however, that a function *is* uniquely defined over all *positive* time by its Laplace transform. The formula for the inverse Laplace transform requires the idea of contour integration which may not be sufficiently familiar to you or indeed known to you at all at the present time. Accordingly we will not discuss this topic in much detail.

Inverse Laplace Transform

Given $F(s)$ the inverse Laplace transform is defined as the following integral:

$$f(t) = \mathcal{L}^{-1}(F(s)) = \frac{1}{2\pi j} \int_{\Gamma} F(s) e^{st} \, ds$$

where Γ is a line in the complex plane parallel to the imaginary axis and having the property that all of the *singularities* of the complex function $F(s)$ lie to the left of this line. The evaluation of the integral is achieved by applying Cauchy's residue theorem.

Inverse Laplace Transform

Mercifully in engineering we very rarely have a need for this rather sophisticated definition. Indeed so rare is the need that it is probably fair to say that the majority of engineers happily go through their careers without ever needing to know the definition. I can say this with some confidence because most engineers that I have met do not even know the names for all of the possible types of singularity that can occur in a complex function. They know of only one, the *pole*.

Inverse Laplace Transform

In fact, in engineering when it comes to looking for the inverse Laplace transform, almost invariably we are looking for the inverse Laplace transform of a rational polynomial and in this case we may employ far less sophisticated techniques. The key is a particular type of expansion of a rational polynomial, known as the *partial fraction expansion*.

Partial Fraction Expansion

Assume $F(s)=\dfrac{N(s)}{D(s)}$, N and D polynomials having no common zeros such that $\deg(N)<\deg(D)$ and such that denominator D is monic. Then there exist constants

$$A_{11},\ldots,A_{1m_1},\ldots,A_{p1},\ldots,A_{pm_p}$$

such that

$$F(s)=\frac{A_{11}}{(s-s_1)}+\cdots+\frac{A_{1m_1}}{(s-s_1)^{m_1}}+\cdots+\frac{A_{p1}}{(s-s_p)}+\cdots+\frac{A_{pm_p}}{(s-s_p)^{m_p}}$$

Note

A *monic* polynomial is a polynomial having unity as the coefficient of the highest power.

The resulting expansion of $F(s)$ is called the *partial fraction expansion*. The values of s for which $F(s)$ is infinity (i.e. $s_1, \dots , s_p$) are called the *poles* of F. Pole s_i is said to be of *order* m_i. A pole of order one is said to be *simple*. A pole of order more than one is said to a *multiple* pole.

Inverse Laplace Transform

Assume

$$F(s)=\frac{A_{11}}{(s-s_1)}+\cdots+\frac{A_{1m_1}}{(s-s_1)^{m_1}}+\cdots+\frac{A_{p1}}{(s-s_p)}+\cdots+\frac{A_{pm_p}}{(s-s_p)^{m_p}}$$

then

$$f(t)=A_{11}\exp(s_1t)u(t)+\cdots+A_{1m_1}\frac{t^{m_1-1}}{(m_1-1)!}\exp(s_1t)u(t)$$

$$+\cdots+A_{p1}\exp(s_pt)u(t)+\cdots+A_{pm_p}\frac{t^{m_p-1}}{(m_p-1)!}\exp(s_pt)u(t)$$

is the inverse Laplace transform of $F(s)$.

Proof

The proof is based on the fact that

$$\mathcal{L}\left(\frac{t^n}{n!}e^{-\alpha t}u(t)\right)=\frac{1}{(s+\alpha)^{n+1}}$$

This follows from the shifting property and from the fact that:

$$\mathcal{L}\left(\frac{t^n}{n!}u(t)\right)=\frac{1}{s^{n+1}}$$

To establish this note that: $\mathcal{L}(u(t))=\dfrac{1}{s}$

Proof

Now employ the differentiation property

$$\mathcal{L}\left(-tf\left(t\right)\right)=\frac{dF\left(s\right)}{ds}$$

$$\mathcal{L}\left(-tu\left(t\right)\right)=\frac{d}{ds}\left(\frac{1}{s}\right)=-\frac{1}{s^{2}} \qquad \mathcal{L}\left(tu\left(t\right)\right)=\frac{1}{s^{2}}$$

Now make the inductive assumption:

$$\mathcal{L}\left(\frac{t^{n}}{n!}u\left(t\right)\right)=\frac{1}{s^{n+1}}$$

Proof

Now employ the differentiation property again

$$\mathcal{L}\left(-tf(t)\right) = \frac{dF(s)}{ds}$$

$$\mathcal{L}\left(-t\left(\frac{t^n}{n!}u(t)\right)\right) = \frac{d}{ds}\left(\frac{1}{s^{n+1}}\right) = -\frac{n+1}{s^{n+2}}$$

$$\mathcal{L}\left(\frac{t^{n+1}}{(n+1)!}u(t)\right) = \frac{1}{s^{n+2}}$$

The result follows by induction.

Proof

So $$\mathcal{L}\left(\frac{t^n}{n!}e^{-\alpha t}u(t)\right)=\frac{1}{(s+\alpha)^{n+1}}$$

and of course $\frac{t^n}{n!}e^{-\alpha t}u(t)$ is zero for $t<0$. We may deduce that

$$\frac{t^n}{n!}e^{-\alpha t}u(t)=\mathcal{L}^{-1}\left(\frac{1}{(s+\alpha)^{n+1}}\right)$$

A typical term in the partial fraction expansion has the form $\frac{A_{ik}}{(s-s_i)^k}$

$$\mathcal{L}^{-1}\left(\frac{A_{ik}}{(s-s_i)^k}\right)=\frac{t^{k-1}}{(k-1)!}e^{s_i t}u(t)$$

Proof

By setting $\alpha = -s_i$ and $n = k$-1 we obtain

$$\mathcal{L}^{-1}\left(\frac{A_{ik}}{\left(s-s_i\right)^k}\right)=\frac{t^{k-1}}{\left(k-1\right)!}e^{s_i t}u\left(t\right)$$

The proof is completed by observing a consequence of the linearity property of the Laplace transform, namely that the inverse Laplace transform of a sum is the sum of the inverse Laplace transforms.

Note

The purpose of evaluating the partial fraction expansion of a rational polynomial in the context of Laplace transform theory is that it permits one to determine the inverse Laplace transform by inspection. You do need to practice a little bit, but in principle this statement is true.

Example 1.3

$$F(s) = \frac{s^2 - 6}{s^3 + 4s^2 + 3s}$$

$$= \frac{5/2}{(s+1)} + \frac{1/2}{(s+3)} - \frac{2}{s}$$

$$f(t) = \left(\tfrac{5}{2}\exp(-t) + \tfrac{1}{2}\exp(-3t) - 2\right)u(t)$$

Code

The objective of this section is to learn how to solve linear, constant coefficient ordinary differential equations with given initial conditions. The primary tool which will be available to you to assist with this endeavour is the computer package Matlab. As our first application of Matlab let us see how to solve the previous inverse Laplace transform problem by employing the package.

Matlab Command Line

I will present Matlab code in a style of my own. In this style >> denotes the command prompt which you will see in the command window when you start up Matlab. The colour code which I adopt is blue for once-off commands, black for repeated commands, *green italic* for comments and red for other. Mostly we will be employing once-off commands. Matlab is an interpreted C language. This means that if you enter a legal once-off command Matlab and then hit the return key Matlab goes off, compiles the command, executes it and presents the results, usually almost immediately.

Example 1.3

For the problem at hand the key Matlab command which we require is the **residue** command. Matlab command names generally are all lower case letters. Almost all Matlab commands require an input argument or arguments. These follow the name of the command in parenthesis (i.e. round brackets) separated by commas if there are more than one. To get information on a Matlab command, once we know the name of the command we type help followed by the command name. For example:

```
>> help residue
```

Example 1.3

Obviously this method of getting help only works when you already know the command name. It remains useful however because often in Matlab one command name refers to a suite of commands, with the input arguments specifying which of this suite of commands you actually wish to employ. It can be hard to remember them all, so help is actually helpful. For our problem the help file for the residue command indicates that we will need to give two input arguments, the first will give information about the numerator polynomial and the second information about the denominator polynomial of the rational polynomial $F(s)$. The best method is to create Matlab polynomials equal to these numerator and denominator polynomials.

Example 1.3

$$F(s) = \frac{s^2 - 6}{s^3 + 4s^2 + 3s}$$

```
>>  Num = [1 0 -6]      create numerator polynomial Num
>>  Den = [1 4 3 0]     create denominator polynomial Den
```

A polynomial in *s* is created in Matlab by creating a vector whose elements are the coefficients in descending powers of *s*. Note *all* coefficients must be present, even if they are zero. A vector is created by subtending a set of numbers in square brackets with the individual numbers separated by a space.

Example 1.3

It is reasonable to question (as the more astute may already have done) how Matlab can tell the difference between a vector and a polynomial. The answer is that it cannot. It only learns that you mean it to interpret the vector as the coefficients of a polynomial when this vector is employed as the input argument of a command which would require that this argument is a polynomial. For example the **roots** command will calculate the roots of a polynomial. It requires one input argument which will actually be a vector, but it will interpret the numbers as describing a polynomial.

Example 1.3

$$F(s)=\frac{s^2-6}{s^3+4s^2+3s}$$

```
>>  roots(Num)       find the roots of polynomial Num

>>  roots(Den)       find the roots of polynomial Den
```

Now for the present example we wish, in preparation for finding the inverse Laplace transform, to find the partial fraction expansion, and this is what the residue command does.

```
>>  [R,P,K] = residue(Num,Den)       find the partial fraction
expansion of F
```

Example 1.3

Note that I have employed the residue command by using output arguments, R, P and K, as well as input arguments. If I employ the residue command without output arguments I get this:

```
>>  residue(Num,Den)          no output arguments
```

```
ans =

    0.5000
    2.5000
   -2.0000
```

The command outputs as an answer three numbers.

Example 1.3

Comparing with the actual answer I see that these numbers appear as the coefficients. So the residue command without output arguments tells me the coefficients (or numerators) of the partial fraction expansion. It tells me nothing about the denominators however, and does not clearly inform me as to which coefficient goes with which denominator.

$$F(s)=\frac{s^2-6}{s^3+4s^2+3s}=\frac{5/2}{(s+1)}+\frac{1/2}{(s+3)}-\frac{2}{s}$$

Example 1.3

On the other hand, employing the residue command with three output arguments gives:

```
>> [R,P,K] = residue(Num,Den)
```

output arguments

```
R =

    0.5000
    2.5000
   -2.0000


P =

    -3
    -1
     0


K =

     []
```

We get two sets of three numbers. The first set being again the coefficients, the second set giving the denominators with the ordering of these sets being such that the first belongs to the first *etc*.

Example 1.3

There is a last output, namely K = [ ]. This involves a class of problems which we will not be considering so the simplest thing to do is to just ignore it. We interpret the outputs R and P as informing us that the partial fraction expansion is:

$$F(s)=\frac{R(1)}{s-P(1)}+\frac{R(2)}{s-P(2)}+\frac{R(3)}{s-P(3)}=\frac{1/2}{(s+3)}+\frac{5/2}{(s+1)}-\frac{2}{s}$$

We get the same partial fraction expansion as above, although granted not in the same order. R by the way stands for residues and P for poles.

Example 1.4

$$F(s) = \frac{5s^3 - 6s - 3}{s^3(s+1)^2}$$

$$= \frac{3}{s} - \frac{3}{s^3} - \frac{3}{(s+1)} + \frac{2}{(s+1)^2}$$

$$f(t) = \left(3 - \tfrac{3}{2}t^2 - 3\exp(-t) + 2t\exp(-t)\right)u(t)$$

Example 1.4

$$F(s) = \frac{5s^3 - 6s - 3}{s^3(s+1)^2}$$

```
>> Num = [5 0 -6 -3]     create numerator polynomial Num
>> Den = [1 2 1 0 0 0]   create denominator polynomial Den
```

Again the first step towards finding the inverse Laplace transform is to find the partial fraction expansion. This problem is more difficult than the last because the poles are multiple, i.e. the denominator polynomial has a triple root at 0 and a double root at -1.

Example 1.4

$$F(s) = \frac{5s^3 - 6s - 3}{s^3(s+1)^2}$$

>> [R,P,K] = residue(Num,Den) *find the partial fraction expansion of F*

R =

-3
2
3
0
-3

P =

-1
-1
0
0
0

The command also gives K = [ ] which I ignore.

Example 1.4

Many students have great difficulty at first in interpreting the data R and P in the case where there are multiple poles. Matlab is actually saying in this case that the partial fraction expansion is:

$$F(s) = \frac{R(1)}{s-P(1)} + \frac{R(2)}{(s-P(2))^2} + \frac{R(3)}{s-P(3)} + \frac{R(4)}{(s-P(4))^2} + \frac{R(5)}{(s-P(5))^3}$$

$$= \frac{-3}{(s+1)} + \frac{2}{(s+1)^2} + \frac{3}{s} + \frac{0}{s^2} - \frac{3}{s^3}$$

i.e. when successive “poles” (P-values) are equal the denominator changes form, being raised to higher and higher powers.

Example 1.5

Everything is now in place. We may start to solve linear constant coefficient ordinary differential equations:

$$\dot{x} + 2x = 3t + 1 \quad \text{for } t > 0 \qquad\qquad x(0) = 1 \; .$$

Step 1: apply Laplace transform to both sides, employing the properties of linearity and differentiation and employing some known Laplace transforms

$$(s+2)X - 1 = \frac{3}{s^2} + \frac{1}{s}$$

Example 1.5

Step 2: solve the resulting algebraic equation in the *s*-domain.

$$X = \frac{1}{(s+2)} + \frac{3}{s^2(s+2)} + \frac{1}{s(s+2)}$$

Advice: do not try to simplify the resulting expression by combining terms.

Step 3: Find the partial fraction expansion. In this case the partial fraction expansion of the sum of the three components is the sum of the partial fraction expansions of the individual components.

Example 1.5

The first term is trivial, it is already in the form of a partial fraction expansion.

$$\frac{3}{s^2(s+2)} = \frac{0.75}{s+2} - \frac{0.75}{s} + \frac{1.5}{s^2}$$

```
>>  Num = [3]      create numerator polynomial Num
>>  Den = [1 2 0 0]     create denominator polynomial Den
>> [R,P,K] = residue(Num,Den)
```

```
R =                    P =

    0.7500               -2
   -0.7500                0
    1.5000                0
```

Example 1.5

$$\frac{1}{s(s+2)} = -\frac{0.5}{s+2} + \frac{0.5}{s}$$

```
>>  Num = [1]      create numerator polynomial Num
>>  Den = [1 2 0]  create denominator polynomial Den
>> [R,P,K] = residue(Num,Den)
```

```
R =                 P =

  -0.5000             -2
   0.5000              0
```

Example 1.5

Combine and now simplify:

$$X = \frac{1}{(s+2)} + \frac{0.75}{(s+2)} - \frac{0.75}{s} + \frac{1.5}{s^2} - \frac{0.5}{(s+2)} + \frac{0.5}{s}$$

$$= \frac{1.25}{s+2} - \frac{0.25}{s} + \frac{1.5}{s^2}$$

Step 4: Determine $x(t)$ by inspection.

$$x(t) = 1.25e^{-2t} - 0.25 + 1.5t \quad \text{for } t > 0$$

Example 1.5

Step 5: Check

$$x(t) = 1.25e^{-2t} - 0.25 + 1.5t \qquad \text{for } t > 0$$

$$x(0) = 1.25 - 0.25 = 1$$

$$\dot{x} + 2x = \left(-2.5e^{-2t} + 1.5\right) + 2\left(1.25e^{-2t} - 0.25 + 1.5t\right)$$
$$= 1 + 3t$$

The purported solution does indeed satisfy the differential equation and the initial condition.

Example 1.6

Recall the model obtained for the simple gravity pendulum:

$$ml\ddot{\theta} + kl\dot{\theta} + mg\sin(\theta) = 0$$

$$\theta(0) = \theta_0 \quad , \quad \dot{\theta}(0) = 0.$$

This differential equation is nonlinear. There is no input with which we can drive the system to a desired operating point. However, there is a solution for which the system has a preference. Left to its own devices the pendulum will come to rest hanging straight down.

Example 1.6

This special preferred solution is $\theta(t) = 0$ which obviously satisfies the equation:

$$ml\ddot{\theta} + kl\dot{\theta} + mg\sin(\theta) = 0\,,$$

although of course it does not satisfy the given initial conditions. A solution like this, where the system comes to rest is called an *equilibrium*. Absent an input we generally adopt for our operating point conditions the requirement that the system remains close to equilibrium.

Example 1.6

We then look for the equation which describes the time evolution of the offset of the angle from equilibrium subject to the assumption that this offset is small. To this end we note from the Maclaurin series that

$$\sin(\theta) = \theta + \cdots$$

and deduce that the first approximation (linearisation here) is:

$$ml\ddot{\theta} + kl\dot{\theta} + mg\theta = 0\,.$$

Example 1.6

We are faced with a linear, constant-coefficient ODE:

$$ml\ddot{\theta} + kl\dot{\theta} + mg\theta = 0$$

$$\theta(0) = \theta_0 \quad , \quad \dot{\theta}(0) = 0.$$

To be more precise assume the bobbin to have a mass of 1 kg, the length of the pendulum to be 2 m, the damping coefficient to be 0.05 and take g = 9.8 m/sec^2.

$$\ddot{\theta} + 0.05\dot{\theta} + 4.9\theta = 0$$

$$\theta(0) = \theta_0 \quad , \quad \dot{\theta}(0) = 0.$$

Example 1.6

Step 1: apply Laplace transform to both sides, employing the properties of linearity and differentiation

$$s^2\Theta - s\theta_0 + 0.05(s\Theta - \theta_0) + 4.9\Theta = 0.$$

Step 2: solve the resulting algebraic equation in the *s*-domain.

$$\Theta = \frac{(s+0.05)\theta_0}{s^2 + 0.05s + 4.9}.$$

Step 3: Find the partial fraction expansion. In this case it will be θ_0 times a partial fraction expansion.

Example 1.6

$$\frac{(s+0.05)}{s^2+0.05s+4.9}=\frac{0.5-0.0056j}{s+0.025-2.2135j}+\frac{0.5+0.0056j}{s+0.025+2.2135j}$$

```
>>  Num = [1 0.05]          create numerator polynomial Num
>>  Den = [1 0.05 4.9]      create denominator polynomial Den
>> [R,P,K] = residue(Num,Den)
```

```
R =                                  P =

   0.5000 - 0.0056i                   -0.0250 + 2.2135i
   0.5000 + 0.0056i                   -0.0250 - 2.2135i
```

Example 1.6

The situation is not as we have seen before. The residues (i.e. the coefficients) and the poles are not real. Matlab of course uses the symbol i to denote the square root of -1. Our proof of the inverse Laplace transform formula did not make any mention of whether the parameters involved were real. It applies verbatim when they are not.

Step 4: Determine $\theta(t)$ by inspection.

$$\theta(t) = \theta_0(0.5 - 0.0056\,j)e^{-(0.025-2.2135\,j)t} + \theta_0(0.5 + 0.0056\,j)e^{-(0.025+2.2135\,j)t}$$

for t > 0.

Example 1.6

Although this is naturally how the method gives the solution there is an obvious problem with it. It appears to suggest that the solution is complex. Appearances in this case are deceiving.

$$0.5 - 0.0056\,j = Ae^{j\phi}$$

```
>>  A = abs(R(1))        R(1) actually equals 0.5-0.0056j from above
>>  Phi = angle(R(1))
```

A =

0.5000

Phi =

-0.0113

Example 1.6

$$\theta(t) = \theta_0 A e^{j\phi} e^{-0.025t} e^{2.2135\,jt} + \theta_0 A e^{-j\phi} e^{-0.025t} e^{-2.2135\,jt}$$

$$\theta(t) = \theta_0 A e^{-0.025t} \left(e^{j(2.2135t+\phi)} + e^{-j(2.2135t+\phi)} \right)$$

$$\theta(t) = 2\theta_0 A e^{-0.025t} \cos(2.2135t + \phi)$$

$$\theta(t) = \theta_0 (1.0001) e^{-0.025t} \cos(2.2135t - 0.0113)$$

Example 1.6

So apparently the angle is not complex after all. What has happened is that we had two complex functions which were complex conjugates of one another. When we add them the imaginary parts cancel producing a real signal.

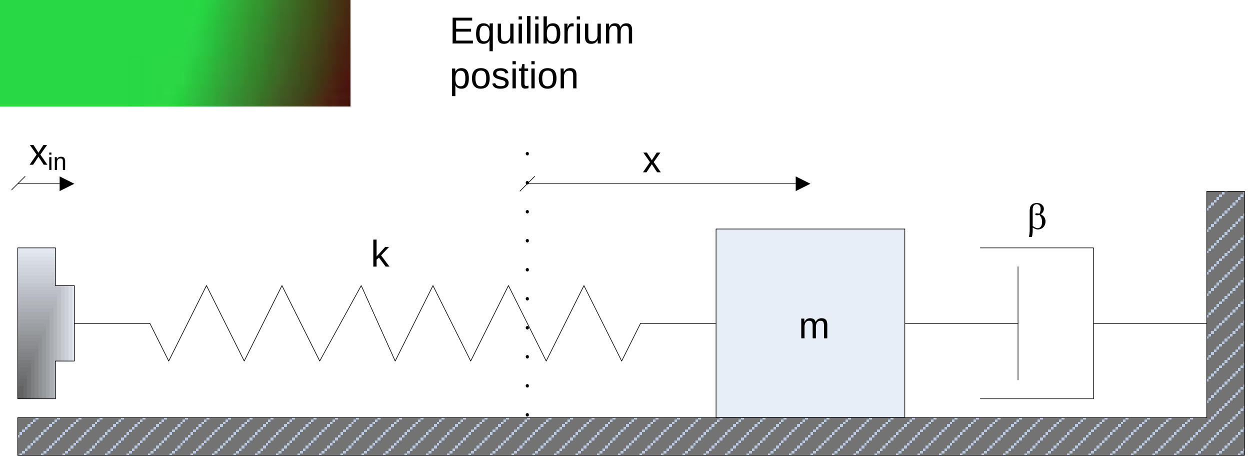
$$\theta(t) = \theta_0 (1.0001) e^{-0.025t} \cos(2.2135t - 0.0113)$$

The formula predicts that the solution will amount to an oscillation at a fixed frequency of 2.2135 rad/sec, with an amplitude which is exponentially decaying with time, albeit somewhat slowly.

Example 1.6

This is a good numerical approximation of the response and accords somewhat with our intuition. It is however quite inaccurate in a number of respects. In fact the frequency of the oscillation does not remain constant, it increases as the amplitude decreases. Moreover in real pendula the amplitude does not merely decay exponentially to zero, it actually becomes zero in finite time. Failure to predict these qualitative properties of the solution is unfortunately the price paid for linearising. The gain of course is a formula which is otherwise unavailable.

Example 1.7:

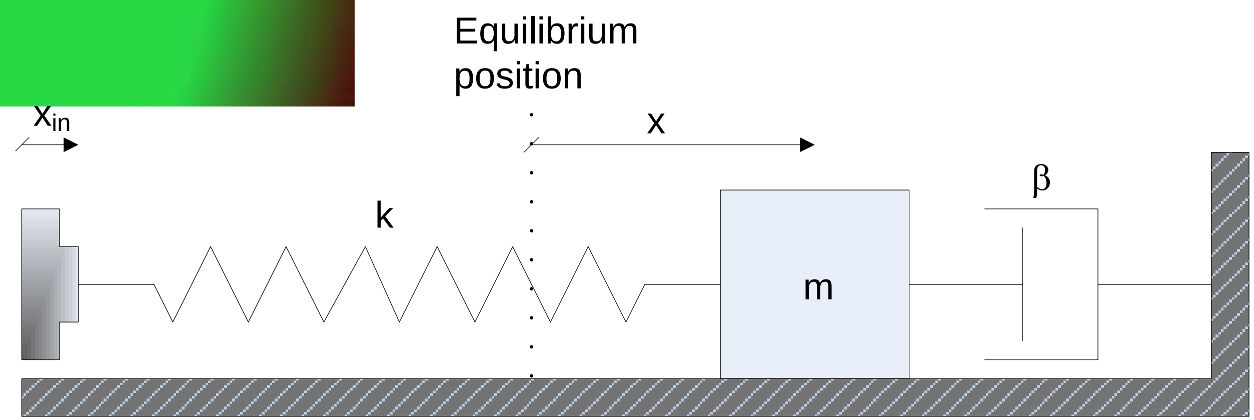


We adopt a standard “linear from the outset” model.

$$m\ddot{x} = -\beta\dot{x} - k\left(x - x_{in}\right)$$

let $m = 2\,\text{kg}$, $k = 8\,\text{N/m}$, $\beta = 1\,\text{Nsec/m}$

Example 1.7:

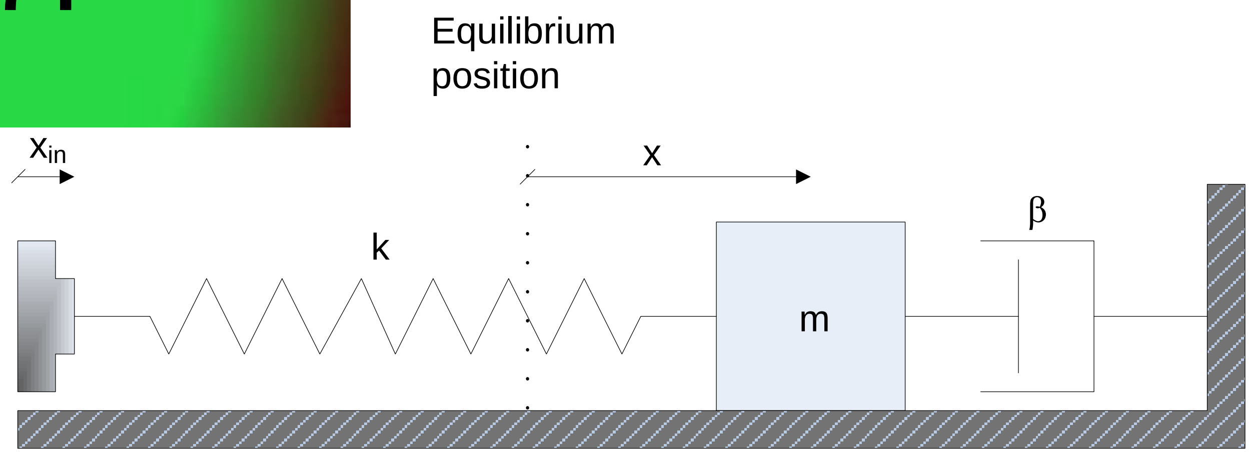


$$2\ddot{x}+\dot{x}+8x=8x_{in}$$

Step 1: apply Laplace transform to both sides, employing the properties of linearity and differentiation

$$2\left(s^2X-sx_0-\dot{x}_0\right)+\left(sX-x_0\right)+8X=8X_{in}$$

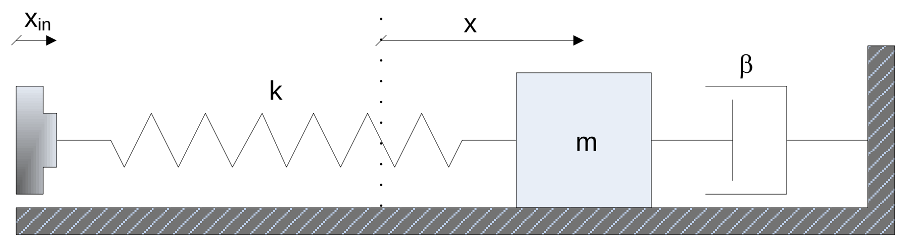
Example 1.7:



Step 2: solve the resulting algebraic equation in the *s*-domain.

$$X(s) = \frac{(2s+1)x_0 + 2\dot{x}_0}{2s^2 + s + 8} + \frac{8X_{in}(s)}{2s^2 + s + 8}$$

Example 1.7



This step is incomplete until I choose the input function. Suppose I consider the response under the conditions where the actuator suddenly moves forward by one unit, i.e. let $x_{in}(t) = u(t)$ giving:

$$X(s) = \frac{(2s+1)x_0}{2s^2+s+8} + \frac{2\dot{x}_0}{2s^2+s+8} + \frac{8}{s(2s^2+s+8)}$$

Step 3: find the partial fraction expansion.

Example 1.7

$$\frac{(s+0.5)}{s^2+0.5s+4}=\frac{0.5-0.063j}{s+0.25-1.9483j}+\frac{0.5+0.063j}{s+0.25+1.9483j}$$

```
>>  Num = [1 0.5]          create numerator polynomial Num
>>  Den = [1 0.5 4]    create denominator polynomial Den
>> [R,P,K] = residue(Num,Den)
```

```
R =

   0.5000 - 0.0630i
   0.5000 + 0.0630i
```

```
P =

  -0.2500 + 1.9843i
  -0.2500 - 1.9843i
```

Example 1.7

$$\frac{1}{s^2+0.5s+4}=\frac{-0.252j}{s+0.25-1.9483j}+\frac{0.252j}{s+0.25+1.9483j}$$

```
>> Num = [1]    create numerator polynomial Num
>> Den = [1 0.5 4]    create denominator polynomial Den
>> [R,P,K] = residue(Num,Den)
```

```
R =

   0 - 0.2520i
   0 + 0.2520i

P =

  -0.2500 + 1.9843i
  -0.2500 - 1.9843i
```

Example 1.7

$$\frac{4}{s\left(s^2+0.5s+4\right)}=\frac{-0.5+0.063j}{s+0.25-1.9483j}+\frac{-0.5-0.063j}{s+0.25+1.9483j}+\frac{1}{s}$$

```
>>  Num = [4]     create numerator polynomial Num
>>  Den = [1 0.5 4 0]     create denominator polynomial Den
>> [R,P,K] = residue(Num,Den)
```

R =	P =
-0.5000 + 0.0630i	-0.2500 + 1.9843i
-0.5000 - 0.0630i	-0.2500 - 1.9843i
1.0000	0

Example 1.7

Combine and now simplify:

$$X = \frac{(0.5-0.063j)x_0 - (0.252j)\dot{x}_0 + (-0.5+0.063j)}{(s+0.25-1.9483j)} +$$
$$\frac{(0.5+0.063j)x_0 + (0.252j)\dot{x}_0 + (-0.5-0.063j)}{(s+0.25+1.9483j)} + \frac{1}{s}$$

Step 4: Determine $x(t)$ by inspection.

Example 1.7

$$x(t)=$$

$$(0.5-0.063j)x_0e^{-(0.25-1.9483j)t}-(0.252j)\dot{x}_0e^{-(0.25-1.9483j)t}$$

$$+(-0.5+0.063j)e^{-(0.25-1.9483j)t}+(0.5+0.063j)x_0e^{-(0.25+1.9483j)t}$$

$$+(0.252j)\dot{x}_0e^{-(0.25+1.9483j)t}+(-0.5-0.063j)e^{-(0.25+1.9483j)t}+1$$

for $t>0$.

Again the method has offered a complex solution when in fact the solution is real. Again this is because we have sums of complex conjugates, so the imaginary parts will cancel. The proper solution is real so we need to affect those cancellations.

Example 1.6

Occurring a great deal is the complex number.

$$z = 0.5 + 0.063j = Ae^{j\phi}$$

```
>> z = 0.5+0.063i
>> A = abs(z)
>> Phi = angle(z)
```

A =

0.5040

Phi =

0.1253

Example 1.7

$$x(t) = Ax_0 e^{-0.25t} e^{-j(1.9483t+\phi)} + Ax_0 e^{-0.25t} e^{j(1.9483t+\phi)}$$
$$-0.252\dot{x}_0 e^{-0.25t} e^{j\left(1.9483t+\frac{\pi}{2}\right)} - 0.252\dot{x}_0 e^{-0.25t} e^{-j\left(1.9483t+\frac{\pi}{2}\right)}$$
$$-Ae^{-0.25t} e^{j(1.9483t-\phi)} - Ae^{-0.25t} e^{-j(1.9483t-\phi)} + 1 \quad \text{for } t > 0.$$

$$x(t) =$$
$$1 + 1.0079x_0 e^{-0.25t}\cos(1.9483t + 0.1253)$$
$$-0.504\dot{x}_0 e^{-0.25t}\cos\left(1.9483t + \tfrac{\pi}{2}\right)$$
$$-1.0079e^{-0.25t}\cos(1.9483t - 0.1253) \quad \text{for } t > 0.$$

Example 1.7

The response is made up of three components. The component:

$$1.0079 x_0 e^{-0.25t} \cos(1.9483t + 0.1253)$$

is due to the initial potential energy shaking out. The component

$$-0.504 \dot{x}_0 e^{-0.25t} \cos\left(1.9483t + \tfrac{\pi}{2}\right)$$

amounts to the initial kinetic energy shaking out. The component

$$1 - 1.0079 e^{-0.25t} \cos(1.9483t - 0.1253)$$

is the forced response. It includes the *steady-state* or long term solution which is 1.

Note

There are a few technical consequences of the differentiation property of the Laplace transform which I will deal with now. Although they appear to be rather useless they are in fact surprisingly important, a fact which we will consider at a later point.

Final Value Theorem

If $\mathcal{L}\left(f\left(t\right)\right)=F\left(s\right)$ then

$$\lim_{s\to 0} sF\left(s\right)=\lim_{t\to\infty} f\left(t\right)$$

Initial Value Theorem

If $\mathcal{L}(f(t)) = F(s)$ then

$$\lim_{s\to\infty} sF(s) = \lim_{t\to 0^+} f(t)$$

Example 1.8

Signal	Laplace Transform
$u(t)$	$\frac{1}{s}$
$\exp(-\alpha t)u(t)$	$\frac{1}{(s+\alpha)}$
$tu(t)=\int_{0^-}^{t} u(\tau)d\tau$	$\frac{1}{s^2}$
$t^2 u(t)=\int_{0^-}^{t} 2\tau u(\tau)d\tau$	$\frac{2}{s^3}$
$t^n u(t)$	$\frac{n!}{s^{n+1}}$
$\frac{t^n \exp(-\alpha t)}{n!}u(t)$	$\frac{1}{(s+\alpha)^{n+1}}$

Example 1.8

Signal	Laplace Transform
$\sin(\omega_0 t)u(t)$	$\dfrac{\omega_0}{s^2+\omega_0^2}$
$\cos(\omega_0 t)u(t)$	$\dfrac{s}{s^2+\omega_0^2}$
$t\sin(\omega_0 t)u(t)$	$\dfrac{2\omega_0 s}{\left(s^2+\omega_0^2\right)^2}$
$t\cos(\omega_0 t)u(t)$	$\dfrac{s^2-\omega_0^2}{\left(s^2+\omega_0^2\right)^2}$

Interpretation

A signal (or function) $f(t)$ is described as being in the *time-domain*. Its Laplace transform $F(s)$ is described as being in the *s-domain*. If one knows the time-domain signal $f(t)$ then it is often relatively simple to find its Laplace transform, i.e. to determine what the signal looks like in the s-domain. We have seen in a few examples that the opposite is also true. If we know the s-domain signal $F(s)$ then it is possible by employing the *inverse Laplace transform* to determine what the signal looks like in the time-domain.

Interpretation

Since we have two “domains” in which we can view the signal and since we can relatively easily flip back and forth between these domains and between these descriptions of the signal a philosophical question may be raised: which domain and which description is the right one? That is not the same question as asking which is the more familiar, the answer to this is plain. Within mathematics a pragmatic and rather general answer has been offered for all questions of this type. If one has two valid descriptions of the one thing, and if one has an effective procedure for jumping back and forth between these descriptions, then neither description is *the* right one, they are both equally correct.

Interpretation

So I contend that the description of a sinusoid for example by the formula:

$$f(t) = 2\cos(3t) \quad , \quad t \geq 0$$

is no more valid and no less valid than the description of this same sinusoid by the formula:

$$F(s) = \frac{2s}{s^2 + 9}$$

Interpretation

The advantage of having multiple equally valid descriptions of an object (a signal here) is that, whereas neither may claim any greater validity than the other, for certain problems one description may result in the problem being easier to solve. Equal validity is not the same as equal utility.

Some of the previous properties may be reconsidered in the light of this statement.

Interpretation

In the time-domain the derivative of signal $f(t)$ may commonly be relatively easily determined, although it can be rather tedious. In the s-domain things are completely different. The differentiation property asserts that:

$$\mathcal{L}\left(\frac{d}{dt}f(t)\right) = sF(s) - f(0).$$

So the act of differentiating in the time domain is mimicked in the s-domain by the act of multiplying by s, with the minor and slightly annoying issue of having to also subtract the initial value.

Interpretation

This reveals the true power of the Laplace transform method. It does not lie in the different descriptions that it offers for signals, rather it lies in the different descriptions which it offers for the effects of mathematical operations upon those signals. So, the annoying subtraction of the initial condition being set aside, it is almost true that the operation of differentiation in the time domain is equal to the much simpler operation of multiplication by s in the s-domain. That is not wonderful, because differentiation is analytically an easy operation. The integration property however shows that integration in the time-domain is equal to division by s in the s-domain and that is wonderful, because integration is analytically very difficult.

Interpretation

The shifting property also shows that the operation of a delay of T sec in the time-domain is equal to the operation of multiplication by $exp(-sT)$ in the s-domain.

So a number of mathematical operations in the time-domain have corresponding operations in the s-domain which amount to (or consistent to a large extent of) multiplying by something, s, $1/s$ or $exp(-sT)$.

Interpretation

That is all very well but it is the likes of example 1.4 that is the crux of the issue:

$$\dot{x} + 2x = f(t) \qquad , \qquad x(0) = 0$$

$$X = \frac{F(s)}{s+2}$$

The mathematical operation of being a solution with zero initial conditions to a forced, linear, constant-coefficient ordinary differential equation in the time-domain also corresponds to multiplication (of the *s*-domain form of the forcing term) by something (1/(*s*+2) here) in the *s*-domain. This is true with great generality.

Section 1 - Conclusion

- The *s*-domain description of many important mathematical operations comprises of multiplying by something.
- Tables of Laplace transforms may be built up by elementary integration and application of the transform properties.
- The Laplace transform forms the basis of a four step systematic process for solving linear, constant coefficient ODEs with given initial conditions.

Section 2 - Outline

- Systems: Linearity, Time-Invariance and Causality.
- Dirichlet's formula.
- Transfer function.

Signals

I have used the phrase “signal” a number of times. That is somewhat sloppy since I have not defined the term. I elect to rectify that situation now.

What do I mean by a signal? The answer is that initially I do not know, but I do know some examples which I feel fit the bill.

Voltage, current, power, pressure, temperature, speed, acceleration, speech, heart beat, ECG, EEG, gross domestic product. These are all quantities regularly dealt with by engineers where I feel that I might reasonably refer to them as signals. These quantities have quite distinct physical natures and as such are very dissimilar

Signals

Mathematics is not about how different things are different. Anyone can see how different things differ. It takes considerably more skill to see how obviously different things are the same. A voltage between the pins of a two pin plug is manifestly different to the speed of a car, so why are they exactly the same?

They are the same because they have some common properties: they usually vary with time and at a particular time they take on a particular value (a real number).

If each of these is to be an example of a *signal* what then is the appropriate definition of a signal?

Signals

Definition 2.1:

A continuous-time signal is a function of a continuous real variable (this variable is commonly given the symbol t denoting time).

A continuous-time signal is commonly called an *analog* (*analogue*) *signal*.

Note

The definition of signal given is by no means the most general possible. Signals do not have to be 1-D (e.g. still images, video), do not have to depend on a single variable (e.g. pressure usually depends on position, x, y, z, and time, t) and do not have to depend on time at all.
In other modules you may take them as functions only of a position variable x for example.

Note

If f is a function of the continuous variable t then the particular number in the range associated with the particular value t in the domain is denoted $f(t)$. We will commonly abuse notation by allowing this same collection of symbols to denote the function itself.

So the voltage between the pins of a two-pin plug and the speed of a car, which are obviously totally different, are, less obviously, exactly the same because they are both examples of continuous-time signals.

Signals

I should note that there is another commonality associated with all of the “signals” on my list above. In each case the signal was associated with a “system”, indeed in a number of cases, ECG and EEG for example, it was obviously an output of a system. Really signals cannot be disconnected from systems, so I had better define what I will mean by a system.

Signals

Informally, a system is a “thing” which has “inputs” and “outputs”, where we often consider that the inputs somehow “cause” the outputs.

An electric fire: input = shaft angle, output = heat.

An internal combustion engine: input = flow of combustible fluid, output = rotational kinetic energy of shaft.

Radio: input = e-m radiation at antenna, output = acoustic wave.

Note

Radio appears to be a poor example of point 3, since e-m radiation at the antenna is not easily adjusted. This is largely because a single radio is not really a system, but rather a part of a much larger system for radio communication. In effect it is a *subsystem*.

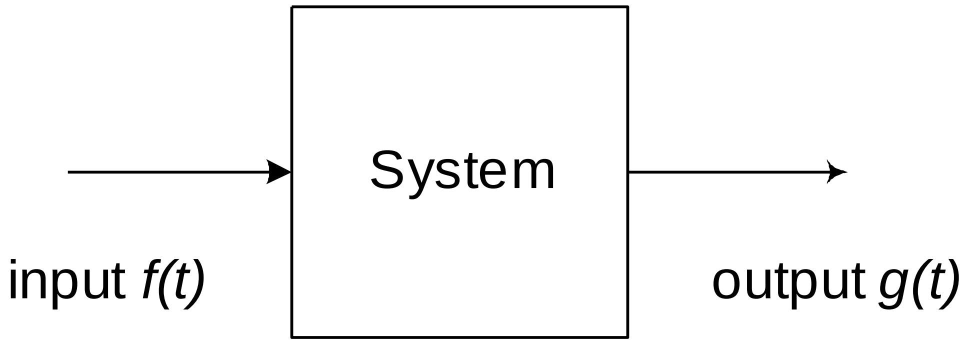
Note

In the examples quoted above the system has one input and one output. Such a system is called a single input/single output (SISO) system.
The inputs and outputs are signals. Their physical natures can be quite distinct.
The physical nature of the input is commonly characterised by its being “readily adjusted”.
The physical nature of the output is commonly “desirable”. It is often why we build the system.

Note

The reason for building a system when the physical natures of input and output are the same is that the manner in which the input is transformed to produce the output may itself be desirable. For example a public broadcast system has acoustic input and output. The purpose of the system of course is that the output is much louder and physically replicated over a large area.

Systems



Definition 2.2: A continuous-time (analogue/analog) single-input single-output (SISO) system processes one continuous-time signal called the input (or excitation) and produces as a result a second continuous-time signal called the output (or response).

Note

In fact this concept of a system is effectively too general. As a scientific civilisation we cannot say much about systems with this degree of generality. We must make some refinements to the definition. Upon making these refinements we will be led to systems which can in general be analysed and which, moreover, comprise such a useful set of models that they will include many if not all of the models that you have seen up to this point in your education. The refinements are two in number: time-invariance and linearity.

Models: Example 3

The liquid level system

$$A\dot{h} = Q_{in} - C_d A_d \sqrt{2gh}$$

may be considered to be a system with input signal equal to the input flow rate Q_{in} and output equal to the liquid level h. We have seen that this is a fairly complicated system from the viewpoint of analysis. We also introduced a simpler related system, namely the system which described how the offset of the level from its rated value evolved with time.

Models: Example 3

This system is: $$\dot{\tilde{h}} + \alpha\tilde{h} = \gamma\tilde{q}_{in}(t) \qquad t \geq 0.$$

We noted as a consequence of linearity that the general solution of this equation equals the sum of a particular solution $\tilde{h}_p(t)$ and a solution $\tilde{h}_0(t)$ to the homogeneous equation

$$\dot{\tilde{h}} + \alpha\tilde{h} = 0 \qquad t \geq 0.$$

I find this unsatisfactory as it has been left since neither the particular solution nor the solution to homogeneous equation are uniquely defined.

Models: Example 3

I noted that an equation such as:

$$\dot{\tilde{h}} + \alpha\tilde{h} = \gamma\tilde{q}_{in}(t) \qquad t \geq 0.$$

is incomplete, describing a mathematical problem which fails to have a unique solution and failing to describe a proper physical problem. To complete the problem statement I must state an initial condition.

$$\dot{\tilde{h}} + \alpha\tilde{h} = \gamma\tilde{q}_{in}(t) \qquad t \geq 0, \quad \tilde{h}(0) = \tilde{h}_0$$

With this addition I may describe a process which uniquely describes both the particular solution and the solution of the associated homogeneous equation.

Models: Example 3

For the particular solution I ask that not only should it be a solution, it should have zero initial condition, i.e.

$$\tilde{h}_p(0) = 0$$

For the solution to the associated homogeneous equation I ask that not only should it be a solution, it should satisfy the given initial condition, i.e.

$$\tilde{h}_0(0) = \tilde{h}_0$$

It follows that the sum $\tilde{h}_0(t) + \tilde{h}_p(t)$ not only satisfies the equation, it also satisfies the initial condition as:

$$\tilde{h}(0) = \tilde{h}_0(0) + \tilde{h}_p(0) = \tilde{h}_0 + 0 = \tilde{h}_0 \; .$$

Models: Example 3

In general the particular solution and the solution to the homogeneous equation are uniquely defined by this process. When this particular choice is made for the particular solution it is commonly called the *forced response*, since it is entirely determined by the forcing term. The unique solution to the associated homogeneous equation is, again in this event, called the *free response*, since it is completely determined by the differential equation and the initial state and completely independent of the forcing term.

Models: Example 3

Physically the forced response corresponds to that component of the response (i.e. the level of the liquid) which is a reaction to the forcing term. The free response on the other hand corresponds to the second (and only other) component of the response which comprises the initial energy stored in the system "shaking itself out".

Models: Example 3

If we focus our attention upon the forced response (the erstwhile particular solution now rendered unique by the additional zero initial condition requirement imposed upon it) we may observe two significant system properties. Firstly, if $\tilde{h}_p(t)$ is the forced response of the system for the input $\tilde{q}_{in}(t)$ then the forced response for input $\sigma\tilde{q}_{in}(t)$ is $\sigma\tilde{h}_p(t)$ for any scalar σ. This property is sometimes known as *homogeneity* although it should be called homogeneity of degree one. I capture its essence thus: “double the input gives double the output”.

Models: Example 3

Secondly, if $\tilde{h}_{p1}(t)$ is the forced response of the system for the input $\tilde{q}_{in,1}(t)$ and if $\tilde{h}_{p2}(t)$ is the forced response of the system for the input $\tilde{q}_{in,2}(t)$ then the forced response for input $\tilde{h}_{p1}(t)+\tilde{h}_{p2}(t)$ is:

$$\tilde{q}_{in,1}(t)+\tilde{q}_{in,2}(t)$$

This property is commonly known as *superposition*. I might summarise it as “the response to the sum is the sum of the responses”. Between them these two properties characterise a linear system.

Time-Invariance

Definition 2.3: A continuous-time, SISO system is time-invariant if it obeys the following rule:

If signal $f(t)$ as input produces signal $g(t)$ as output then delayed signal $f(t-\tau)$ as input produces delayed signal $g(t-\tau)$ as output for any delay τ.

Note

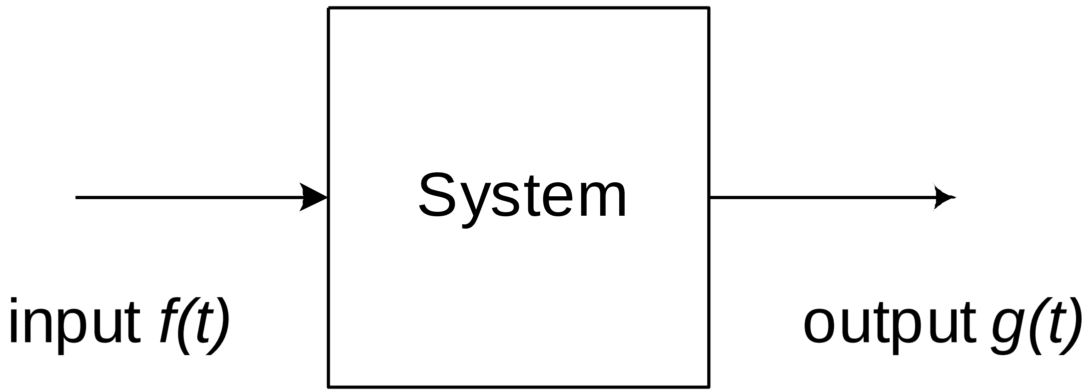
A time-invariant system behaves today in the same way as it behaved yesterday and in the same way as it will behave tomorrow. Obviously all systems decay and no system has existed forever, so strictly speaking time-invariant systems do not exist. However, many systems do have a behaviour which is fairly constant over the useful life of the system and at least over one continuous period of use of the system (*e.g.* a journey). These systems can be approximated by time-invariant systems which are easier to analyse.

Note

A car provides a good example of an approximately time-invariant system. The steering system should behave next time we drive in much the same way as it behaved last time. Driving would be nearly impossible for a normal human being if the system failed to have this approximate time-invariance property. Racing car drivers and military pilots are exceptions in this respect.

Time-Invariance

To illustrate:



Superposition

- One of the most important concepts in classical mathematical physics is Daniel Bernoulli's *Principle of Superposition*.
- e.g. the acceleration of a rigid body due to a sum of applied forces equals the sum of the accelerations due to the individual forces acting alone.

$$\ddot{x} = \frac{\vec{F}_1}{m} \quad , \quad \ddot{x} = \frac{\vec{F}_2}{m}$$

$$\ddot{x} = \frac{\vec{F}}{m} = \frac{\vec{F}_1 + \vec{F}_2}{m} = \frac{\vec{F}_1}{m} + \frac{\vec{F}_2}{m}$$

- An example of the principle in action is the trajectory of a cannon ball. The motion of the ball is a sum of two distinct and completely independent motions: an un-accelerated horizontal motion at the initial horizontal velocity and a vertical motion, downward accelerated by the gravitational force. In this case the principle can be applied and greatly simplifies the problem of finding the trajectory by means of divide and conquer. One seeks the independent motions (vertical and horizontal). These motions are much easier to find (the horizontal motion particularly so). The actual motion is simply the sum.
- This divide and conquer approach, expressing the solution as a sum of simpler components, lies at the very heart of this module.

- Regarding the action of a force as the input and the acceleration as the output then, using system terminology, the superposition principle states that the output of a system to a sum of inputs equals the sum of the outputs to the individual inputs acting alone.

- Informally, a linear system is a system which obeys this more general form of the Principle of Superposition.

Linearity

Definition 2.4: A continuous-time, SISO system is linear if it obeys the following rule:

If signal f_1 as input produces signal g_1 as output and signal f_2 as input produces signal g_2 as output then, given any real numbers α_1 and α_2, composite signal $\alpha_1 f_1 + \alpha_2 f_2$ as input produces composite signal $\alpha_1 g_1 + \alpha_2 g_2$ as output.

e.g. Ideal Resistor (input = current, output = voltage)

e.g. Rigid Body (input = force, output = acceleration)

Note

Linearity implies that if f as input produces g as output, then αf as input produces αg as output for any α. In practice systems are damaged by large input signals so that this relationship fails for large α. Strictly speaking linear systems do not exist. However, many systems do behave approximately linearly provided all signals remain sufficiently small. This observation justifies in part the prevalence of linear system theory, but the main reason is one of mathematical expediency. In general, nonlinear systems defy analysis. Since linear systems are much easier to analyse than nonlinear ones we often accept the inherent inadequacies of the technique and work with linear approximations.

Note

Our refinements to our general definition of a system will be to assume time-invariance and linearity. With these two restrictions upon the systems to be considered we will find that the problem of analysis shifts from one which is unsolved to one which is not only solved, but comprises probably one of the deepest human insights into the workings of the universe. At a minimum we may now reconsider the second method for solving linear, constant-coefficient ODEs considered above.

Interpreting Method of Variation of Parameters

The basis of the Laplace transform method is, I hope, at least somewhat clarified by the previous comments and examples. In the introduction we considered a second method of solving linear constant coefficient ODEs. This was to solve the homogeneous equation by essentially guessing the form of the solution and then checking the details and then to employ the method of variation of parameters to find the particular solution (i.e. a particular solution satisfying zero initial conditions). The next task of the module is to offer a more modern clarification of this method.

Interpreting Method of Variation of Parameters

The method of variation of parameters gave solutions having a certain form, namely the sum of a solution to the homogeneous equation satisfying the initial conditions and an integral from 0 to t of an integrand of the form:

$$\int_0^t h(t-\tau) f(\tau) d\tau.$$

where f is the input or forcing term. We may ask whether the solution always takes this form, whether there is any insight into the behaviour of systems at which this is hinting and whether there are more efficient methods for finding both the solution to the homogeneous equation and the function h appearing in this integral.

Interpreting Method of Variation of Parameters

The answer to each of these questions is affirmative and again the Laplace transform plays a pivotal role in providing these answers. Very surprisingly, although we shall not develop the ideas further, the answers to all of these questions draw from ideas developed in the field of number theory of all places.

Example: 1.2 again

In terms of their Laplace transforms, the Laplace transform of the solution equals the Laplace transform of the particular solution plus that of a solution to the homogeneous equation, with the latter reflecting the initial conditions. We may compare the two components:

$$\tilde{H} = \frac{\tilde{h}(0)}{s+\alpha} + \frac{\gamma \tilde{Q}_{in}(s)}{s+\alpha}.$$

The first component (the free response) is like the second (the forced response) but with a forcing term having a Laplace transform equal to a constant.

Ideal Impulse

So the initial condition has the same effect as a forcing term whose Laplace transform is a constant equal to this initial value. A question may be asked concerning what kind of function has a Laplace transform which is a constant. The answer is no function and it rather appears that that should be the end of the matter. There is a scenario however where this idea of imagining initial conditions to be the effect of forcing terms is not an apparently foolish fantasy.

Ideal Impulse

If a ball is at rest and is struck head on by a second ball of equal mass travelling at a certain velocity conservation of linear momentum asserts that both will travel subsequently with a velocity equal to half that of the original upon impact. The velocity of each undergoes a dramatic and sudden change, an increase for one and a decrease for the second. Really the best way to analyse the situation is to simply start the process again after the collision and view the respective velocities as initial conditions for the post-collision objects.

Ideal Impulse

Alternatively in the case of the ball which was originally at rest we may imagine that there is a simply immense force of extremely short, indeed zero duration which instantaneously accelerates it from rest to half the velocity of the other ball just before collision. Such forces are not called forces in mechanics, they are called *impulses*. So the equation of motion of the first ball is:

$$m\dot{v} = I(t) \qquad m\frac{v_0}{2} - 0 = \int_{-\varepsilon}^{\varepsilon} I(t)\,dt$$

Ideal Impulse

Now the impulse $I(t)$ must have some fairly strange properties. For any $\varepsilon > 0$

$$I(t) = 0 \quad \text{for } |t| > \varepsilon$$

$$m\frac{v_0}{2} = \int_{-\varepsilon}^{\varepsilon} I(t)dt = \int_{-\infty}^{\infty} I(t)dt$$

i.e. the impulse must be zero for all t not equal to 0 yet somehow it must be so immense at $t = 0$ that the area under the curve is non-zero.

Ideal Impulse

Alternatively for any $\varepsilon > 0$ and for any $t > 0$ we may solve

$$m\dot{v} = I(t)$$

$$m\frac{v_0}{2} = \int_{-\varepsilon}^{t} I(\tau)d\tau = \int_{0^-}^{t} I(\tau)d\tau$$

Taking Laplace transforms and using the integration property:

$$\frac{mv_0}{2}\frac{1}{s} = \frac{\mathcal{L}(I(t))}{s} \qquad \mathcal{L}(I(t)) = \frac{mv_0}{2}$$

Ideal Impulse

So by a plausibility argument it seems that the ideal impulse should have a Laplace transform which is a constant. Notwithstanding the presence of this concept within mechanics for a very long period it was not until the early twentieth century that the ideal impulse was formally defined and not until the mid-twentieth century that a reasonably complete theory of the set of mathematical objects to which the ideal impulse belongs was developed.

Dirac delta function

Definition 1.2: The Dirac delta function, $\delta(t)$, is a defined by the following two apparently contradictory properties:

$$\delta(t) = 0 \quad \text{for all } t \neq 0$$

$$\int_{-\infty}^{\infty} \delta(t)\,dt = 1$$

Dirac justifies his definition by imagining the delta function emerge through a limiting process, much as we imagine the ideal impulse to emerge.

Dirac delta function

Indeed Dirac imagines the delta function as being the limit as ε approaches zero from above of the non-deal impulses:

$$\delta_{\varepsilon}(t) = 0 \quad \text{for all } |t| > \varepsilon$$

$$\delta_{\varepsilon}(t) = \frac{1}{2\varepsilon} \quad \text{for all } |t| \leq \varepsilon.$$

The delta function is not a function (it does not map from the real numbers into the real numbers). It is in fact a *generalised function* or *distribution*.

Dirac delta function

The theory of distributions was developed by Laurent Schwartz. It essentially introduces the distributions as relating to the functions in the same way that the real numbers relate to the rational numbers, by being the set of all the limits of sequences of functions. Not only is distribution theory decidedly the youngest mathematical theory of which you will hear mention in your engineering modules, it is also one of the more difficult, winning for Schwartz the Field's medal in 1950.

Dirac delta function

The Laplace transform of the Dirac delta function likewise is defined as the limit as ε approaches zero from above of the Laplace transforms of the associated non-deal impulses:

$$\mathcal{L}(\delta(t)) = \lim_{\varepsilon \to 0^+} \mathcal{L}(\delta_\varepsilon(t)) = \lim_{\varepsilon \to 0^+} \int_{-\varepsilon}^{\varepsilon} \frac{1}{2\varepsilon} e^{-st} dt$$

$$\lim_{\varepsilon \to 0^+} \frac{1}{2\varepsilon} \left(\frac{e^{-s\varepsilon} - e^{s\varepsilon}}{-s} \right) = \lim_{\varepsilon \to 0^+} \frac{1}{2\varepsilon} \left(\frac{1 - s\varepsilon - 1 - s\varepsilon}{-s} \right) = 1$$

Example: 1.2 again

$$\tilde{H} = \frac{\tilde{h}(0)}{s+\alpha} + \frac{\gamma \tilde{Q}_{in}(s)}{s+\alpha}.$$

So the response of the liquid-level system may equally be regarded as the response of a system having zero initial conditions to an input or forcing term:

$$\left(\tilde{h}(0) \big/ \gamma \right) \delta(t) + \tilde{q}_{in}(t).$$

This observation is the first which only partially generalises, but it comprises a vehicle for introducing the delta function.

Dirac delta function: Sifting

The Dirac delta function has a very important property known as the *sifting* property. Given any suitable "test function" φ

$$\int_{-\infty}^{\infty} \delta(t-t_0)\varphi(t)dt = \varphi(t_0)$$

So when you see the Dirac delta function appear in an integral it may look very complicated, but in fact the integral is very easy to evaluate, you just evaluate the function multiplying the delta function at the value of the variable for which the argument of the delta function is zero.

Dirac delta function: Sifting

As it is important let us “prove” this property:

$$\int_{-\infty}^{\infty} \delta(t-t_0)\varphi(t)dt = \int_{-\infty}^{t_0-\varepsilon} \delta(t-t_0)\varphi(t)dt + \int_{t_0-\varepsilon}^{t_0+\varepsilon} \delta(t-t_0)\varphi(t)dt + \int_{t_0+\varepsilon}^{\infty} \delta(t-t_0)\varphi(t)dt$$

We now employ Cauchy’s marvellous *first mean value theorem for integration*:

$$\int_{-\infty}^{\infty} \delta(t)\varphi(t)dt = \varphi(a)\int_{t_0-\varepsilon}^{t_0+\varepsilon} \delta(t-t_0)dt \qquad \text{some } a \in [t_0-\varepsilon, t_0+\varepsilon].$$

Dirac delta function: Sifting

Now note:

$$1 = \int_{-\infty}^{\infty} \delta(t)\,dt = \int_{-\infty}^{\infty} \delta(t - t_0)\,dt =$$

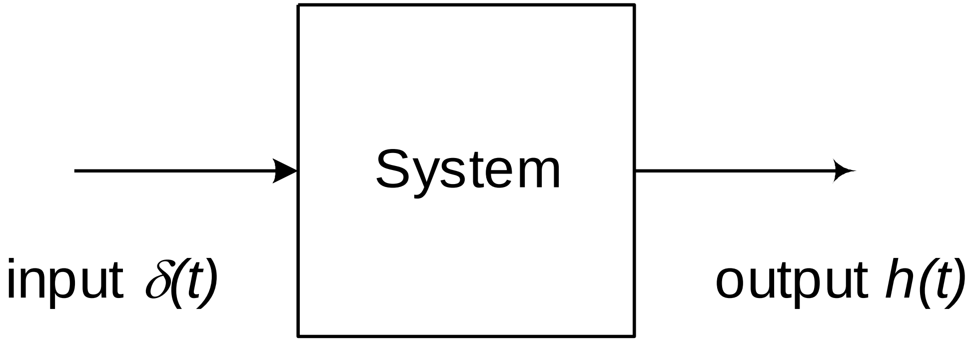
$$\int_{-\infty}^{t_0-\varepsilon} \delta(t - t_0)\,dt + \int_{t_0-\varepsilon}^{t_0+\varepsilon} \delta(t - t_0)\,dt + \int_{t_0+\varepsilon}^{\infty} \delta(t - t_0)\,dt = \int_{t_0-\varepsilon}^{t_0+\varepsilon} \delta(t - t_0)\,dt$$

$$\int_{-\infty}^{\infty} \delta(t)\varphi(t)\,dt = \varphi(a) \qquad \text{some } a \in [t_0 - \varepsilon, t_0 + \varepsilon]$$

and this is true apparently for all $\varepsilon > 0$. By taking the limit as ε approaches 0 from above (assuming the limit exists) we obtain the sifting property

Impulse Response

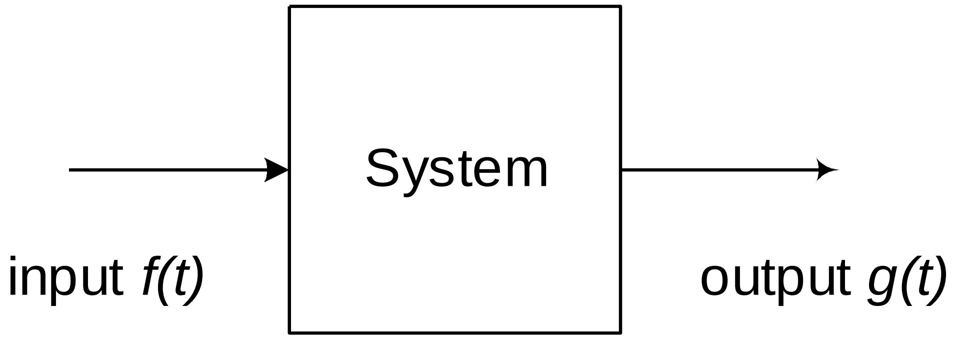
Definition 2.5: The response of a SISO, linear time-invariant (LTI) system to an input equal to the delta function $\delta(t)$ is called the *impulse response* of the system.



The Convolution Integral

Theorem 1: Let the impulse response of a SISO, LTI system be $h(t)$. Then, under certain conditions, the response satisfying zero initial conditions $g(t)$ to an excitation $f(t)$ is given by Dirichlet's formula

$$g(t) = \int_{0^-}^{t} f(\tau) h(t-\tau) d\tau$$



Note

This is a major result of this module. We saw integrals of this kind emerge when we considered to method of variation of parameters in the examples of the introductory lectures. More generally this theorem is explaining how that particular integral for the forced response comes about. Moreover it offers a method for determining the function $h(t)$, since it observes that this function equals the impulse response of the system. Actually we do not consider this method for finding $h(t)$ to be effective.

Argument

i/p $\delta(t) \Rightarrow$ o/p $h(t)$.

i/p $\delta(t-\tau) \Rightarrow$ o/p $h(t-\tau)$ (time-invariance).

i/p $f(\tau)\delta(t-\tau) \Rightarrow$ o/p $f(\tau)h(t-\tau)$ (linearity).

i/p $\int_{0^-}^{t} f(\tau)\delta(t-\tau)d\tau \Rightarrow$ o/p $\int_{0^-}^{t} f(\tau)h(t-\tau)d\tau$ (linearity?).

i/p $f(t) \Rightarrow$ o/p $\int_{0^-}^{t} f(\tau)h(t-\tau)d\tau$ (sifting).

Example

$$\text{Impulse response } h(t) = \begin{cases} \exp(-t) & \text{if } t \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\text{Input } f(t) = \begin{cases} \cos(t) & \text{if } t \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\text{Response } g(t) = \begin{cases} \dfrac{\cos\left(t - \pi/4\right)}{\sqrt{2}} - \dfrac{\exp(-t)}{4} & \text{if } t \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

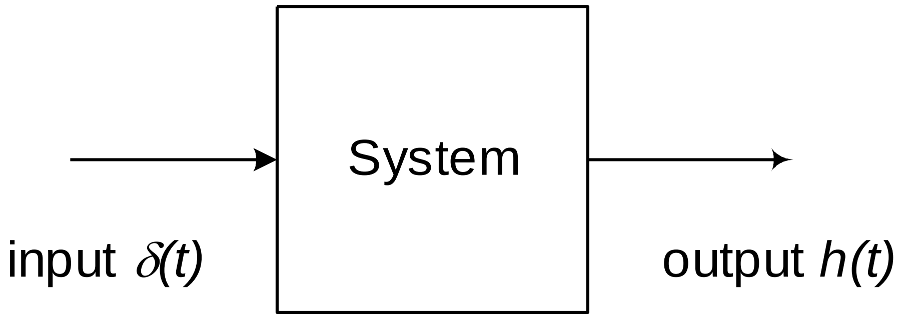
Note

We see that the application of Dirichlet's formula is non-trivial. As we cannot change the formula itself (since it is derived directly from the definition of linearity and time-invariance) we must try to change our perspective in order to achieve some degree of simplification. The fact that the Laplace transform method was also able to solve general equations gives a hint that perhaps the Laplace transform may have something useful to say. In fact we derive a property which in many ways encompasses all of the previous properties.

Causal Systems

Definition 2.6: A continuous time signal $f(t)$ is *causal* if $f(t) = 0$ for all $t < 0$.

Definition 2.7: An LTI, SISO system is causal if the impulse response $h(t)$ is causal.



Note

A non-causal system must have an impulse response which is non-zero for some time $t < 0$. Since the impulse itself happens at the time $t = 0$ the system is therefore anticipating that an impulse is going to happen *before* it actually does happen. As this amounts to unusual behaviour we may have confidence that real systems will indeed be causal.

Laplace Transform: Properties 6

Theorem 2: Convolution.

Let two signals be causal and suppose that

$$\mathcal{L}\left(f_1(t)\right) = F_1(s) \text{ and } \mathcal{L}\left(f_2(t)\right) = F_2(s)$$

then

$$\mathcal{L}\left(\int_{0^-}^{t} f_1(\tau) f_2(t-\tau)\,d\tau\right) = F_1(s)F_2(s)$$

Note

The time integral above is called the convolution of signals f_1 and f_2 and is denoted $f_1 * f_2(t)$. We note how similar this horrible convolution integral is to the integral appearing in Dirichlet's formula.

The Transfer Function

Theorem 3: Let the impulse response of a SISO, LTI, causal system be $h(t)$, let the input $f(t)$ be a causal signal then the output $g(t)$ is causal and

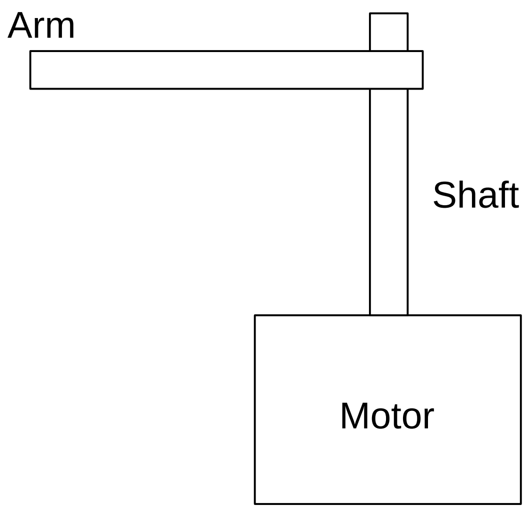
$$G(s) = F(s)H(s)$$

Note

From the correct perspective all LTI, SISO systems look like multipliers. The function $H(s)$ is called the *transfer function*. It equals the Laplace transform of the impulse response. It also equals the ratio of the Laplace transform of the forced response (i.e. the particular response satisfying zero initial conditions) to the Laplace transform of the input.

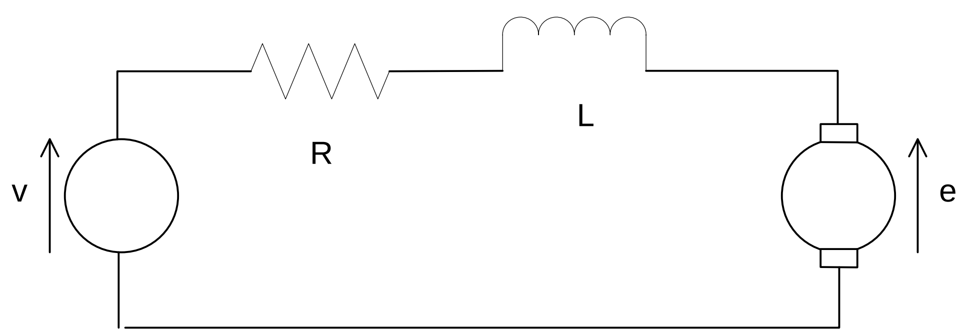
Example 2.1

A robotic arm is fixed to the shaft of a DC motor and free to rotate in the horizontal plane. We control the input voltage to the motor.



Example 2.1

We have to first model the system. Obviously a first attempt is to employ standard models. On the electrical side one of the simplest standard models of a DC motor is as shown. The physical effects modelled are the resistance of the stator wire, the self-inductance of this wire and the back emf generated by the rotor.



$$e = k_e \frac{d\theta}{dt}$$

Example 2.1

Of course this electrical model follows the “linear from the outset approach” which I have previously disparaged. I must bear in mind that appropriate values for the various parameters should be chosen for when the motor is close to its desired operating point, including for example its desired temperature. I must recognise that should the device depart significantly from these desired settings then the chosen parameter values will no longer be appropriate and indeed the model may fail. An obvious failure mechanism for example would be saturation of the iron in the core.

$$v = Ri + L\frac{di}{dt} + k_e\frac{d\theta}{dt}$$

Example 2.1

On the mechanical side we employ substantially the rotational form of Newton's law (sometimes called Euler's law).

$$J\frac{d^2\theta}{dt^2} = T = T_m - k_f\frac{d\theta}{dt}$$

Here J denotes the moment of inertia of the arm. T is the torque (or moment) applied. This is comprised of two components: the *motor torque* T_m and the *friction torque* which we assume proportional to angular velocity. For the motor torque we have the important relationship:

$$T_m = k_e i$$

Example 2.1

We obtain as a result the following set of equations:

$$v = Ri + L\frac{di}{dt} + k_e\frac{d\theta}{dt} \qquad J\frac{d^2\theta}{dt^2} = k_e i - k_f\frac{d\theta}{dt}$$

$$i = \frac{J}{k_e}\frac{d^2\theta}{dt^2} + \frac{k_f}{k_e}\frac{d\theta}{dt}$$

$$v = \frac{RJ}{k_e}\frac{d^2\theta}{dt^2} + \frac{Rk_f}{k_e}\frac{d\theta}{dt} + \frac{LJ}{k_e}\frac{d^3\theta}{dt^3} + \frac{Lk_f}{k_e}\frac{d^2\theta}{dt^2} + k_e\frac{d\theta}{dt}$$

Example 2.1

Or rearranging:

$$\frac{d^3\theta}{dt^3}+\left(\frac{R}{L}+\frac{k_f}{J}\right)\frac{d^2\theta}{dt^2}+\left(\frac{Rk_f}{LJ}+\frac{k_e^2}{L}\right)\frac{d\theta}{dt}=\frac{k_e}{L}v$$

We obtain a linear, constant-coefficient ODE with forcing term $v(t)$, being the voltage input to the motor and response $\theta(t)$ being the angle of the arm. We may take the Laplace transform of both sides and apply the transform properties, specifically differentiation and linearity.

Example 2.1

$$\left(s^3\Theta - s^2\theta(0) - s\dot{\theta}(0) - \ddot{\theta}(0)\right) + \left(\frac{R}{L} + \frac{k_f}{J}\right)\left(s^2\Theta - s\theta(0) - \dot{\theta}(0)\right)$$

$$+\left(\frac{Rk_f}{LJ} + \frac{k_e^2}{L}\right)\left(s\Theta - \theta(0)\right) = \frac{k_e}{L}V(s)$$

$$\Theta = \frac{\frac{k_e}{L}V(s)}{s^3 + \left(\frac{R}{L} + \frac{k_f}{J}\right)s^2 + \left(\frac{Rk_f}{LJ} + \frac{k_e^2}{L}\right)s} +$$

$$\frac{\left(s^2\theta(0) + s\dot{\theta}(0) + \ddot{\theta}(0)\right) + \left(\frac{R}{L} + \frac{k_f}{J}\right)\left(s\theta(0) + \dot{\theta}(0)\right) + \left(\frac{Rk_f}{LJ} + \frac{k_e^2}{L}\right)\left(\theta(0)\right)}{s^3 + \left(\frac{R}{L} + \frac{k_f}{J}\right)s^2 + \left(\frac{Rk_f}{LJ} + \frac{k_e^2}{L}\right)s}$$

Example 2.1

$$\Theta = \frac{N(s)}{D(s)}V(s) + \frac{Q(s)}{D(s)} = H(s)V(s) + \frac{Q(s)}{D(s)} =$$

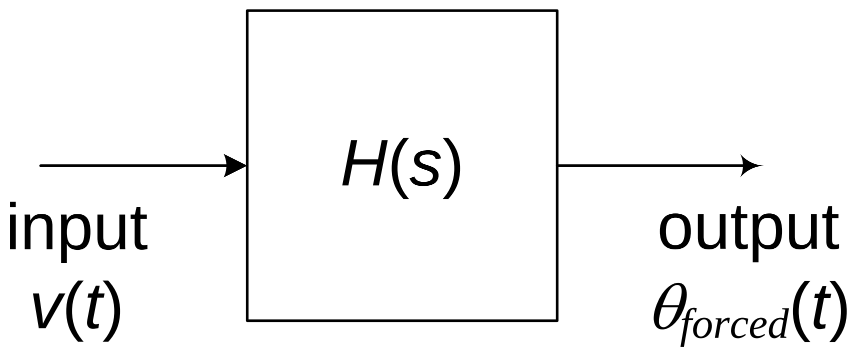
$$N(s) = \frac{k_e}{L} \quad , \quad D(s) = s^3 + \left(\frac{R}{L} + \frac{k_f}{J}\right)s^2 + \left(\frac{Rk_f}{LJ} + \frac{k_e^2}{L}\right)s$$

$$Q(s) =$$

$$\left(s^2\theta(0) + s\dot{\theta}(0) + \ddot{\theta}(0)\right) + \left(\frac{R}{L} + \frac{k_f}{J}\right)\left(s\theta(0) + \dot{\theta}(0)\right) + \left(\frac{Rk_f}{LJ} + \frac{k_e^2}{L}\right)\left(\theta(0)\right)$$

Example 2.1

$$H(s) = \frac{\frac{k_e}{L}}{s^3 + \left(\frac{R}{L} + \frac{k_f}{J}\right)s^2 + \left(\frac{Rk_f}{LJ} + \frac{k_e^2}{L}\right)s}$$



s-domain or *block diagram* model of DC motor and arm

In general: Solving linear, constant coefficient ODEs

$$a_n \frac{d^n}{dt^n} g(t) + \cdots + a_1 \frac{d}{dt} g(t) + a_0 g(t)$$

$$= b_{n-1} \frac{d^{n-1}}{dt^{n-1}} f(t) + \cdots + b_1 \frac{d}{dt} f(t) + b_0 f(t)$$

where $f(t)$ is a specified causal function and initial conditions $g(0^-), g^{(1)}(0^-), \ldots, g^{(n-1)}(0^-)$ are given.

Note

As noted above, many systems are approximately described (at least close to an equilibrium or operating point) by linear, constant coefficient ODEs. In the context of system theory we may view the equation as characterising the performance of a system with inputs f, $g(0^-)$, $g^{(1)}(0^-)$, …, $g^{(n-1)}(0^-)$ and output g. Viewed in this way the equation describes a MISO, LTI system.

Solving linear, constant coefficient ODEs

$$a_n\left(s^n G(s) - s^{n-1}g\left(0^-\right) - s^{n-2}g^{(1)}\left(0^-\right) - \cdots - g^{(n-1)}\left(0^-\right)\right)$$
$$+ \cdots + a_1\left(sG(s) - g\left(0^-\right)\right) + a_0 G(s)$$

$$= b_{n-1}s^{n-1}F(s) + \cdots + b_1 sF(s) + b_0 F(s)$$

$$\left(a_n s^n + \cdots + a_1 s + a_0\right)G(s) = \left(b_{n-1}s^{n-1} + \cdots + b_1 s + b_0\right)F(s)$$
$$+ \left(a_n g^{(n-1)}\left(0^-\right) + \cdots + a_1 g\left(0^-\right)\right) + \cdots + s^{n-1}a_n g\left(0^-\right)$$

Solving linear, constant coefficient ODEs

$$N(s) = b_{n-1}s^{n-1} + \cdots + b_1 s + b_0$$

$$D(s) = a_n s^n + \cdots + a_1 s + a_0$$

$$Q(s) = \left(a_n g^{(n-1)}\left(0^-\right) + \cdots + a_1 g\left(0^-\right)\right) + \cdots + s^{n-1} a_n g\left(0^-\right)$$

$$D(s)G(s) = N(s)F(s) + Q(s)$$

$$G(s) = \frac{N(s)}{D(s)}F(s) + \frac{Q(s)}{D(s)} = H(s)F(s) + \frac{Q(s)}{D(s)}$$

Note

Polynomials N and D are very simply related to the ODE. Polynomial Q depends on the initial conditions and is zero if these are zero. The key component of the Laplace transform method is that we can divide by $D(s)$. The resulting expression for $G(s)$ (and the linearity property of the Laplace transform) indicates that the solution $g(t)$ can be written as a sum of two terms: the forced (particular) solution and the free (homogeneous) solution, where the latter is zero if the initial conditions are zero.

Note

The rational polynomial $H(s)$ is the transfer function of the SISO, LTI system with input $f(t)$ and output equal to the particular solution (i.e. the solution under zero initial conditions). Hence $H(s)$ is the Laplace transform of the impulse response of the system (under zero initial conditions). It is thus that the Laplace transform method offers an efficient procedure for finding the function $h(t)$ appearing in the method of variation of parameters. This function is the impulse response which is the inverse Laplace transform of the transfer function which can be written down by inspection directly from the ODE.

Solving linear, constant coefficient ODEs

$$G(s) = H(s)F(s) + \frac{Q(s)}{D(s)}$$

$$= H(s)F(s) + \frac{A_{11}}{s - s_1} + \cdots + \frac{A_{1m_1}}{(s - s_1)^{m_1}} + \cdots + \frac{A_{p1}}{s - s_p} + \cdots + \frac{A_{pm_p}}{(s - s_p)^{m_p}}$$

$$g(t) = \int_0^t h(t - \tau) f(\tau) d\tau +$$

$$A_{11}e^{s_1 t} + \cdots + A_{1m_1} \frac{t^{m_1 - 1}}{(m_1 - 1)!} e^{s_1 t} + \cdots + A_{p1}e^{s_p t} + \cdots + A_{pm_p} \frac{t^{m_p - 1}}{(m_p - 1)!} e^{s_p t}$$

Note

The Laplace transform method establishes with a high degree of generality certain observations which we made in those special cases where the method of variation of parameters was applied to solve linear, constant coefficient ODEs. The solution to the homogeneous equation is a sum of exponentials whose exponents are the roots of a particular polynomial closely associated with the ODE, namely the polynomial $D(s)$. This solution is a little more complicated than stated when some of these roots are multiple. The forced response on the other hand is given as what we now recognise as a convolution integral, with the critical function $h(t)$ appearing in this integral being the impulse response of the system.

Note

So the Laplace transform method does not just provide an efficient procedure for solving linear, constant coefficient ODEs in the event that the forcing term is such that its Laplace transform is a rational polynomial. It provides a mathematical structure within which we can both explain and interpret other such methods. It has become in effect one of just two methods for solving linear, constant coefficient ODEs about which an engineer will need to know.

Frequency Response

In the introductory examples three broad points were being made.

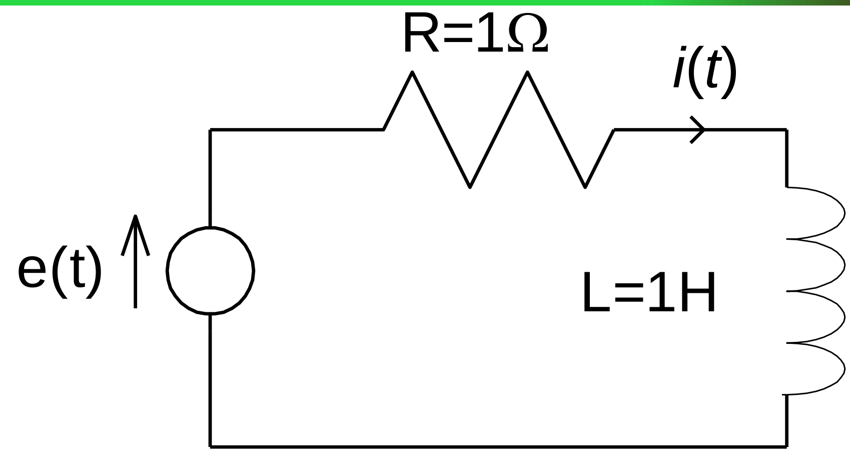
(i) that the Laplace transform method comprised an effective method for solving linear, constant coefficient ODEs. This point has been explored further.

(ii) that the method of variation of parameters offered a solution in the form of the sum of the free response and the forced response. This point has also been explored with the special form of the forced response being explained in terms of convolution and the special form of the free response by the partial fraction expansion and its inverse transform.

Frequency Response

(iii) that the forced response of a linear, constant coefficient ODE to a co-sinusoidal forcing term was a co-sinusoid of the same frequency. This point requires some elaboration. Indeed our further exploration of this point will be extensive. However we start with a simple example.

Example 2.2: Circuits



$$\frac{d}{dt} i(t) + i(t) = e(t), \qquad i(0^-) = I_0$$

$$N(s) = 1, \qquad D(s) = s + 1, \qquad Q(s) = I_0$$

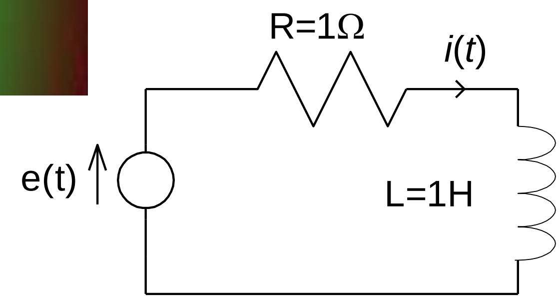
$$H(s) = \frac{1}{s+1}$$

Note

This has the form of a linear, constant coefficient ODE with output (response) $g(t)$ being the current $i(t)$ and input (forcing term) $f(t)$ being the voltage/emf $v(t)$.

First we consider a simple input voltage, the unit step.

Example 2.2



$$I(s) = H(s)E(s) + \frac{Q(s)}{D(s)} = \frac{E(s)}{s+1} + \frac{I_0}{s+1}$$

$$e(t) = u(t), \quad I(s) = \frac{1}{s(s+1)} + \frac{I_0}{s+1}$$

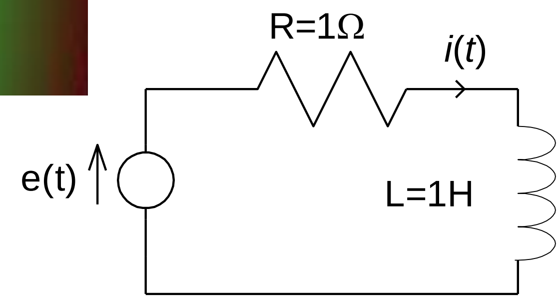
$$I(s) = \frac{(I_0 - 1)}{s+1} + \frac{1}{s}, \qquad i(t) = \left((I_0 - 1)\exp(-t) + 1 \right) u(t)$$

Note

Note that the solution is a sum of two components – a transient which exponentially decays at a rate determined by the poles of the system and a steady state which is independent of the initial conditions. In this case the steady-state is a constant current of 1 A. Of course the input is a constant voltage of 1 V (for $t > 0$) in this case. So DC (i.e. zero frequency) input has produced DC output *in the long run*.

Consider a co-sinusoidal input.

Example 2.2

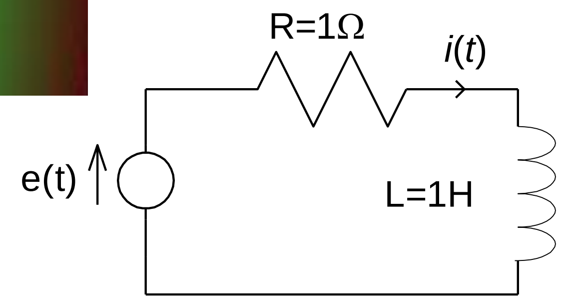


$$I(s) = H(s)E(s) + \frac{Q(s)}{D(s)} = \frac{E(s)}{s+1} + \frac{I_0}{s+1}$$

$$e(t) = 2\cos(t), \quad I(s) = \frac{2s}{(s^2+1)(s+1)} + \frac{I_0}{s+1}$$

$$I(s) = \frac{(I_0 - 1)}{s+1} + \frac{0.5 + j0.5}{s+j} + \frac{0.5 - j0.5}{s-j}$$

Example 2.2



$$I(s) = \frac{(I_0 - 1)}{s+1} + \frac{0.5 + j0.5}{s+j} + \frac{0.5 - j0.5}{s-j}$$

$$i(t) = (I_0 - 1)e^{-t} + (0.5 + j0.5)e^{-jt} + (0.5 - j0.5)e^{jt} \quad \text{for } t > 0$$

$$i(t) = (I_0 - 1)e^{-t} + \cos(t) + \sin(t) \ \text{ for } t > 0$$

$$i(t) = (I_0 - 1)e^{-t} + \sqrt{2}\cos\left(t - \tfrac{\pi}{4}\right) \ \text{ for } t > 0$$

Note

Note that again the solution is a sum of two components –

a transient which exponentially decays at a rate determined by the poles of the system and a steady state which is independent of the initial conditions. In this case the steady-state is a co-sinusoid of frequency 1 rad/sec. Of course the input is a co-sinusoid of frequency 1 rad/sec. So 1 rad/sec input has produced 1 rad/sec output *in the long run*.

Frequency Response

Consider the more general situation where a system is described by a linear, constant-coefficient ODE with input equal to a co-sinusoid

$$f(t) = A\cos(\omega_0 t + \phi)$$

A is the amplitude, ϕ is the phase (in radians) and ω_0 is the frequency (in rad/sec).

Frequency Response

$$G(s)=\frac{N(s)}{D(s)}F(s)+\frac{Q(s)}{D(s)}=H(s)F(s)+\frac{Q(s)}{D(s)}$$

$$f(t)=A\cos(\phi)\cos(\omega_0 t)-A\sin(\phi)\sin(\omega_0 t)$$

$$F(s)=A\cos(\phi)\left(\frac{s}{s^2+\omega_0^2}\right)-A\sin(\phi)\left(\frac{\omega_0}{s^2+\omega_0^2}\right)=\frac{N_F(s)}{s^2+\omega_0^2}$$

Frequency Response

$$G(s) = H(s)A\cos(\phi)\left(\frac{s}{s^2+\omega_0^2}\right) - H(s)A\sin(\phi)\left(\frac{\omega_0}{s^2+\omega_0^2}\right) + \frac{Q(s)}{D(s)}$$

$$G(s) = \frac{A_1}{s+j\omega_0} + \frac{\bar{A}_1}{s-j\omega_0} + \frac{A_{11}}{s-s_1} + \cdots$$

where s_1, ... denote the roots of $D(s)$, i.e. the poles of the transfer function $H(s)$. Assume that all of these poles have negative real part so that all but the first two terms of the partial fraction expansion correspond to transient terms in the solution.

Frequency Response

It is time to discuss Heaviside's method for finding the coefficients of the partial fraction expansion. Note that:

$$G(s)=\frac{N(s)}{D(s)}F(s)+\frac{Q(s)}{D(s)}=\frac{N(s)N_F(s)}{D(s)(s^2+\omega_0^2)}+\frac{Q(s)}{D(s)}=\frac{\tilde{N}(s)}{D(s)(s^2+\omega_0^2)}$$

$$G(s)=\frac{\tilde{N}(s)}{(s^2+\omega_0^2)D(s)}=\frac{\tilde{N}(s)}{(s^2+\omega_0^2)\,(s-s_1)^{m_1}\cdots(s-s_p)^{m_p}}$$

$$=\frac{A_1}{s+j\omega_0}+\frac{A_2}{s-j\omega_0}+\frac{A_{11}}{s-s_1}+\cdots$$

Frequency Response

I have assumed that the poles of $G(s)$ have negative real part. In particular they are not purely imaginary and therefore not equal to +/- $j\omega_0$. Heaviside multiplies both sides of the equation by $s + j\omega_0$.

$$\frac{\tilde{N}(s)}{(s - j\omega_0)\,(s - s_1)^{m_1}\cdots(s - s_p)^{m_p}} = \frac{\tilde{N}(s)}{(s - j\omega_0)\,D(s)}$$

$$A_1 + \frac{A_2(s + j\omega_0)}{s - j\omega_0} + \frac{A_{11}(s + j\omega_0)}{s - s_1} + \cdots$$

Frequency Response

Now Heaviside evaluates both sides of the equation at $s = -j\omega_0$.

$$\frac{\tilde{N}(-j\omega_0)}{(-2j\omega_0)\,D(-j\omega_0)} = A_1$$

$$A_1 = \frac{N(-j\omega_0)N_F(-j\omega_0) + \left((-j\omega_0)^2 + \omega_0^2\right)Q(-j\omega_0)}{(-2j\omega_0)\,D(-j\omega_0)}$$

$$= H(-j\omega_0)\frac{N_F(-j\omega_0)}{(-2j\omega_0)}$$

Frequency Response

In the same manner:

$$A_2 = H(j\omega_0)\frac{N_F(j\omega_0)}{(2j\omega_0)}$$

$$N_F(s) = A\cos(\phi)s - A\sin(\phi)\omega_0$$

$$A_1 = H(-j\omega_0)\frac{(A\cos(\phi)(-j\omega_0) - A\sin(\phi)\omega_0)}{(-2j\omega_0)}$$

$$A_2 = H(j\omega_0)\frac{(A\cos(\phi)(j\omega_0) - A\sin(\phi)\omega_0)}{(2j\omega_0)}$$

Frequency Response

Accordingly:

$$G(s) = \frac{H(-j\omega_0)\dfrac{(A\cos(\phi)(-j\omega_0) - A\sin(\phi)\omega_0)}{(-2j\omega_0)}}{s + j\omega_0} + \frac{H(j\omega_0)\dfrac{(A\cos(\phi)(j\omega_0) - A\sin(\phi)\omega_0)}{(2j\omega_0)}}{s - j\omega_0} + \cdots$$

$$g(t) = H(-j\omega_0)\frac{(A\cos(\phi)(-j\omega_0) - A\sin(\phi)\omega_0)}{(-2j\omega_0)}e^{-j\omega_0 t} + H(j\omega_0)\frac{(A\cos(\phi)(j\omega_0) - A\sin(\phi)\omega_0)}{(2j\omega_0)}e^{j\omega_0 t} + \cdots$$

with all of the omitted terms being transient terms.

Frequency Response

So the steady-state response is:

$$A\cos(\phi)\frac{\left(H(-j\omega_0)e^{-j\omega_0 t}+H(j\omega_0)e^{j\omega_0 t}\right)}{2}+$$

$$A\sin(\phi)\frac{\left(H(-j\omega_0)e^{-j\omega_0 t}-H(j\omega_0)e^{j\omega_0 t}\right)}{2j}$$

Now, H being a real rational polynomial.

$$H(-j\omega_0)=\overline{H}(j\omega_0) \qquad H(j\omega_0)=\left|H(j\omega_0)\right|e^{jArg(H(j\omega_0))}$$

Frequency Response

So the steady-state response is:

$$A\cos(\phi)\left|H(j\omega_0)\right|\frac{\left(e^{-j(\omega_0 t+Arg(H(j\omega_0)))}+e^{j(\omega_0 t+Arg(H(j\omega_0)))}\right)}{2}+$$

$$A\sin(\phi)\left|H(j\omega_0)\right|\frac{\left(e^{-j(\omega_0 t+Arg(H(j\omega_0)))}-e^{j(\omega_0 t+Arg(H(j\omega_0)))}\right)}{2j}$$

$$=A\cos(\phi)\left|H(j\omega_0)\right|\cos(\omega_0 t+Arg(H(j\omega_0)))-$$
$$A\sin(\phi)\left|H(j\omega_0)\right|\sin(\omega_0 t+Arg(H(j\omega_0)))$$

$$=A\left|H(j\omega_0)\right|\cos(\omega_0 t+\phi+Arg(H(j\omega_0)))$$

Frequency Response

So the steady-state response of a system with transfer function $H(s)$ to a co-sinusoidal input

$$f(t) = A\cos(\omega_0 t + \phi)$$

is a co-sinusoidal output of the same frequency as the input:

$$A\left|H(j\omega_0)\right|\cos(\omega_0 t + \phi + Arg(H(j\omega_0)))$$

The output co-sinusoid has a different amplitude and phase.

Frequency Response

The output amplitude is the input amplitude multiplied by:

$$\left|H\left(j\omega_0\right)\right|$$

the modulus of the transfer function evaluated at $s = j\omega_0$. This is called the *gain* at the frequency ω_0.

The output phase is the input phase plus:

$$Arg\left(H\left(j\omega_0\right)\right)$$

the phase or argument of the transfer function evaluated at $s = j\omega_0$.

Frequency Response

The transfer function evaluated at $s = j\omega$ is called the frequency response of the system. It tells us how the system reacts to a sinusoidal input. If the frequency response at a particular frequency ω is rather close to zero, i.e. the modulus is very small, then the system gain at that frequency is very small. Accordingly even a rather large amplitude input at this frequency will produce as a result a very low amplitude output. The system is said to *block* or *filter out* this frequency.

Frequency Response

On the other hand, if the frequency response at a particular frequency ω is rather large, i.e. the modulus is quite large, then the system gain at that frequency is high. Accordingly the system shows a resonance at this frequency since for a modest input signal amplitude at this frequency the output signal amplitude is higher, possibly considerably higher.
I can acquire essentially this same result in a different way, more like the examples of the introduction.

Solving linear, constant coefficient ODEs with sinusoidal forcing term

$$a_n \frac{d^n}{dt^n} g(t) + \cdots + a_1 \frac{d}{dt} g(t) + a_0 g(t)$$

$$= b_{n-1} \frac{d^{n-1}}{dt^{n-1}} f(t) + \cdots + b_1 \frac{d}{dt} f(t) + b_0 f(t).$$

$f(t)$ is a co-sinusoid $cos(\omega t)$ and we seek only the forced response. We may solve by simply guessing that the solution takes the form $Acos(\omega t+\phi)$, although the previous Laplace transform-based method proves this to be commonly the case.

Solving linear, constant coefficient ODEs with sinusoidal forcing term

We wish, substantially to just put the purported solution $A\cos(\omega t+\phi)$ into the equation and check whether there are values for A and ϕ such that it satisfies it. Towards this end it is annoying that the derivatives of this function, and indeed of the forcing function $\cos(\omega t)$ take two forms. The odd order derivatives are sinusoids of radian frequency ω and the even order derivatives are co-sinusoids again of radian frequency ω. As in the special case above we will be faced with an equation involving both co-sinusoids and sinusoids. Within this framework this is a tricky equation to solve. For this reason we make the jump to complex functions.

Solving linear, constant coefficient ODEs with sinusoidal forcing term

$$a_n \frac{d^n}{dt^n} g(t) + \cdots + a_1 \frac{d}{dt} g(t) + a_0 g(t)$$

$$= b_{n-1} \frac{d^{n-1}}{dt^{n-1}} f(t) + \cdots + b_1 \frac{d}{dt} f(t) + b_0 f(t).$$

$$f(t) = \cos(\omega t) = \frac{e^{j\omega t} + e^{-j\omega t}}{2}$$

$$g(t) = A\cos(\omega t + \phi) = \frac{Ae^{j\phi}e^{j\omega t} + Ae^{-j\phi}e^{-j\omega t}}{2}$$

Solving linear, constant coefficient ODEs with sinusoidal forcing term

$$\frac{Ae^{j\phi}}{2}\left(a_n(j\omega)^n+\cdots+a_1(j\omega)+a_0\right)e^{j\omega t}+$$

$$\frac{Ae^{-j\phi}}{2}\left(a_n(-j\omega)^n+\cdots+a_1(-j\omega)+a_0\right)e^{-j\omega t}$$

$$=\frac{1}{2}\left(b_{n-1}(j\omega)^{n-1}+\cdots+b_1(j\omega)+b_0\right)e^{j\omega t}+$$

$$\frac{1}{2}\left(b_{n-1}(-j\omega)^{n-1}+\cdots+b_1(-j\omega)+b_0\right)e^{-j\omega t}.$$

Solving linear, constant coefficient ODEs with sinusoidal forcing term

Again this messy looking equation takes the form

$$z_1 e^{j\omega t} = z_2 e^{-j\omega t}.$$

where z_1 and z_2 are just complex numbers, with z_2 indeed being minus the complex conjugate of z_1. As before we deduce that the only possible solution to these equations is offered by requiring that z_1 and z_2 are zero, with the latter following from the former.

$$z_1 = \frac{Ae^{j\phi}}{2}\left(a_n (j\omega)^n + \cdots + a_1 (j\omega) + a_0\right) - \frac{1}{2}\left(b_{n-1}(j\omega)^{n-1} + \cdots + b_1 (j\omega) + b_0\right) = 0\,.$$

Solving linear, constant coefficient ODEs with sinusoidal forcing term

In terms of the notation introduced above we see a reoccurrence of the polynomial $N(s)$ and $D(s)$. Indeed in terms of these polynomials the previous, apparently tricky equations simplify indeed:

$$Ae^{j\phi} = \frac{N(j\omega)}{D(j\omega)}.$$

Of course $N(s)$ and $D(s)$ are just the numerator and denominator polynomials of the transfer function $H(s)$ of the system. In terms of this transfer function we obtain an even further simplification:

$$Ae^{j\phi} = H(j\omega).$$

Solving linear, constant coefficient ODEs with sinusoidal forcing term

So if the input to a system described by a linear, constant coefficient ODE is a co-sinuosoid $cos(\omega t)$ then the forced response will be the co-sinusoid $A\cos(\omega t+\phi)$. The amplitude and phase of this response are derived from the transfer function evaluated at $s = j\omega$ through the formula:

$$Ae^{j\phi} = H\left(j\omega\right).$$

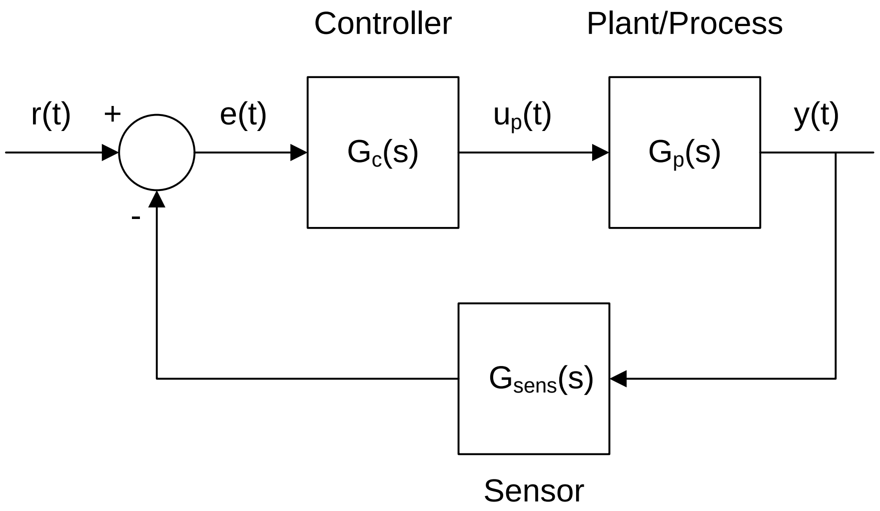
This is substantially the same result as derived above, or rather the special case of it for unit input amplitude and zero input phase. This method however must assume that the solution takes a certain form, the previous method proves it and determines conditions for it to be true.

Solving linear, constant coefficient ODEs with sinusoidal forcing term

Of course we can actually perform the experiment of injecting co-sinusoids (although low amplitude ones) of varying frequency into the system and measuring the resulting response. If this response is also a co-sinusoid of the same frequency as the input then we have evidence that the system is behaving like a linear, time-invariant system. In short apparently we are sufficiently close to the operating point. By finding the ratio of the amplitude of the output to that of the input (we cannot in general inject a signal of unit amplitude since this will be too big. Recall the system is not linear, it only looks linear if we stay close to the operating point) and by finding the phase difference between the output and input we can experimentally determine $H(j\omega)$. This is a powerful method of *system identifica*tion.

Example 2.3

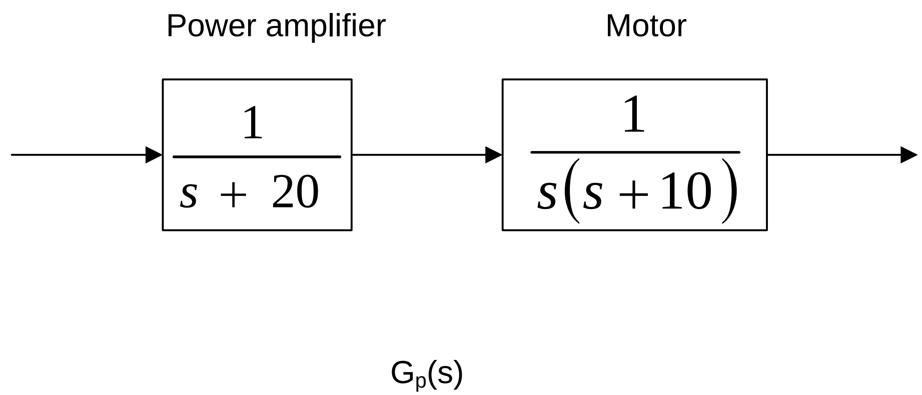
A system to control the position of one of the wheels of a mobile robot is shown. Note that block-diagram models are employed.



The *plant* refers to the actual dynamics of the wheel position.

Example 2.3

A motor positions the wheel, but as it requires some power a power amplifier is also required.



So the plant is actually a cascade of two subsystems. The sensor is assumed to be linear with a unit change in position resulting in a 0.1 V change in sensor output voltage. Accordingly the sensor is modelled as a gain of 0.1.

Example 2.3

For the controller we select a system having transfer function:

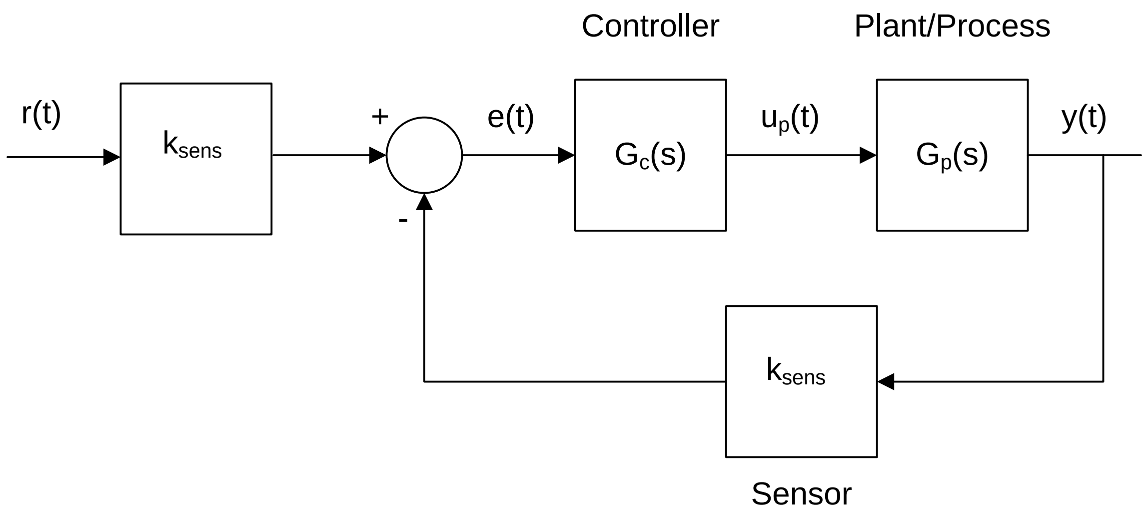
Controller

$$\frac{80{,}000(s+11)}{(s+62)}$$

The choice of the form of this controller and the parameter values employed is the subject of *control theory*.

Example 2.3

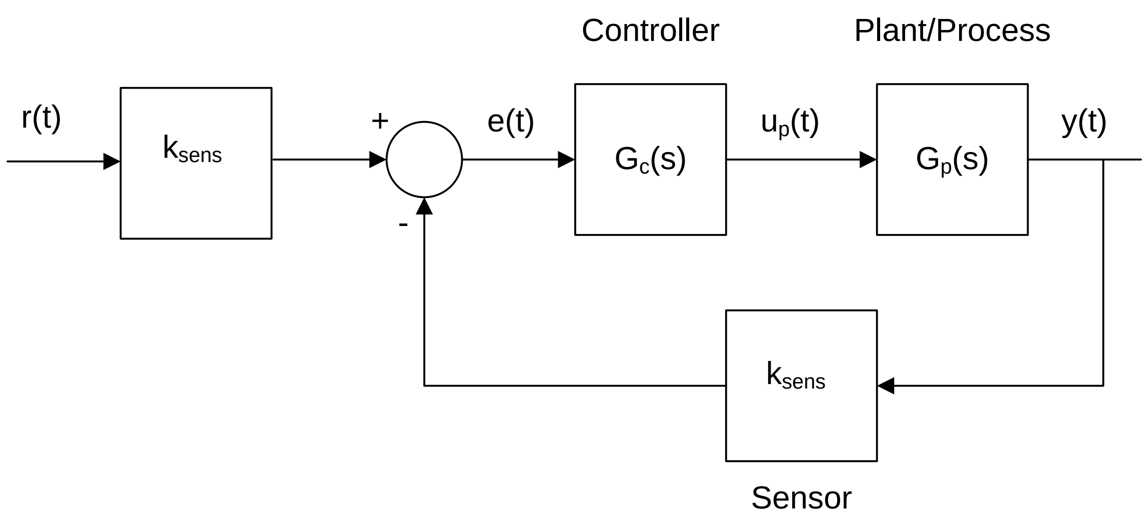
Change input by adding pre-amplifier.



$$Y(s) = G_p(s)G_c(s)\left(k_{sens}R(s) - k_{sens}Y(s)\right)$$

$$Y(s) = \frac{G_p(s)G_c(s)k_{sens}}{1 + G_p(s)G_c(s)k_{sens}}R(s)$$

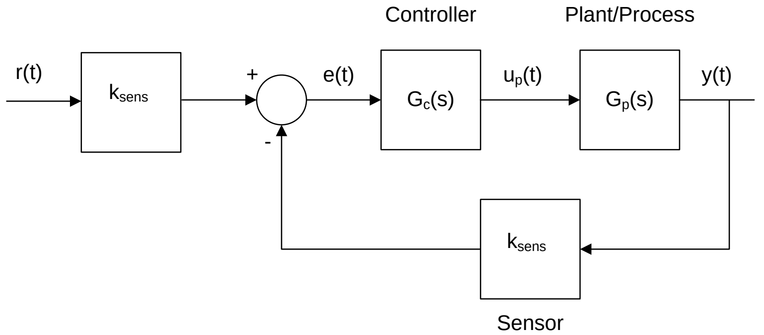
Example 2.3



The transfer function of the system with input $r(t)$ and output $y(t)$ is

$$\frac{G_p(s)G_c(s)k_{sens}}{1+G_p(s)G_c(s)k_{sens}}=\frac{8{,}000s+88{,}000}{s^4+92s^3+2{,}060s^2+20{,}400s+88{,}000}$$

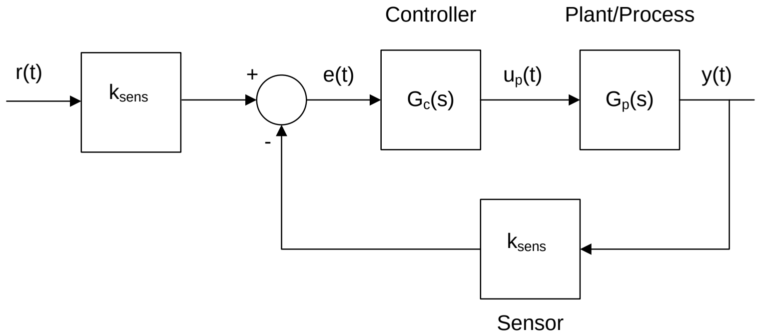
Example 2.3



So making all linearising assumptions from the outset, the system with output $y(t)$ and input $r(t)$ is described by the following linear, constant coefficient ODE provided it remains close to a suitable operating point.

$$\frac{d^4y}{dt^4}+92\frac{d^3y}{dt^3}+2{,}060\frac{d^2y}{dt^2}+20{,}400\frac{dy}{dt}+88{,}000y=8{,}000\frac{dr}{dt}+88{,}000r$$

Example 2.3



In control theory the response of a system to two particular types of input is of great significance. The first is a unit step input, i.e. the response when the input equals the unit step. It is rather problematic even to state this problem in terms of differential equations. In the *s*-domain it is easier however. The unit step response is the response when $R(s) = 1/s$, i.e.

$$Y(s) = \frac{8{,}000s + 88{,}000}{s\left(s^4 + 92s^3 + 2{,}060s^2 + 20{,}400s + 88{,}000\right)}$$

Example 2.3

We employ **residue** to obtain:

$$Y(s)=$$

$$\frac{1}{s}-\frac{0.0381}{s+64.7142}+\frac{0.2195}{s+12.4044}+\frac{(-0.5907+0.5791j)}{s+7.4407-7.3661j}+\frac{(-0.5907-0.5791j)}{s+7.4407+7.3661j}$$

$$y(t)=$$

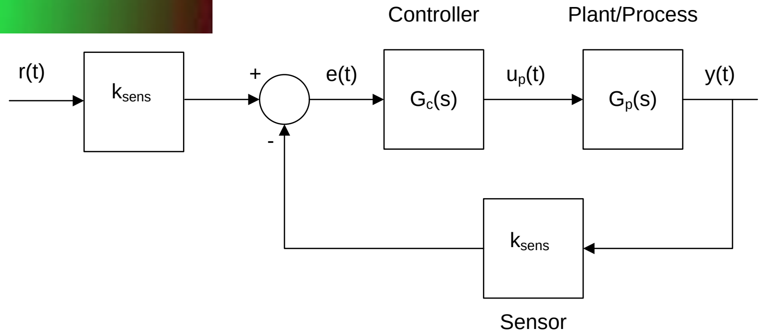
$$1-0.0381e^{-64.7142t}+0.2195e^{-12.4044t}+$$

$$(-0.5907+0.5791j)e^{-7.4407t}e^{j7.3661t}+(-0.5907-0.5791j)e^{-7.4407t}e^{-j7.3661t}$$

$$y(t)=$$

$$1-0.0381e^{-64.7142t}+0.2195e^{-12.4044t}+1.6544e^{-7.4407t}\cos(7.3661t+2.3661)$$

Example 2.3

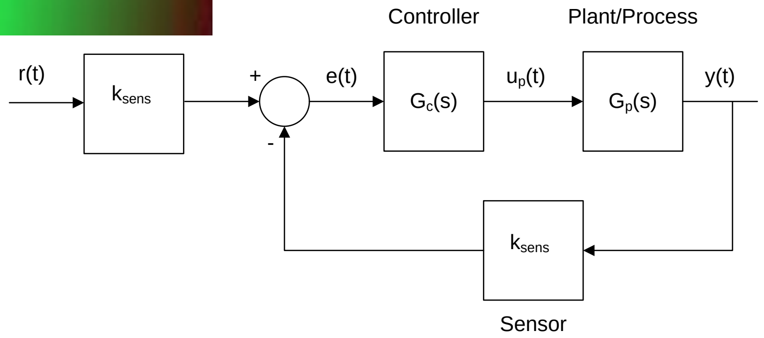


So the long-term or steady-state response of the system to an input which is a unit step is unity. We say that the output *tracks* the input. In fact we could have seen this without taking the inverse Laplace transform by using the final value theorem

$$\lim_{t\to\infty} y(t) = \lim_{s\to 0} sY(s) =$$

$$\lim_{s\to 0} s\left(\frac{8{,}000s + 88{,}000}{s\left(s^4 + 92s^3 + 2{,}060s^2 + 20{,}400s + 88{,}000\right)}\right) = \frac{88{,}000}{88{,}000} = 1$$

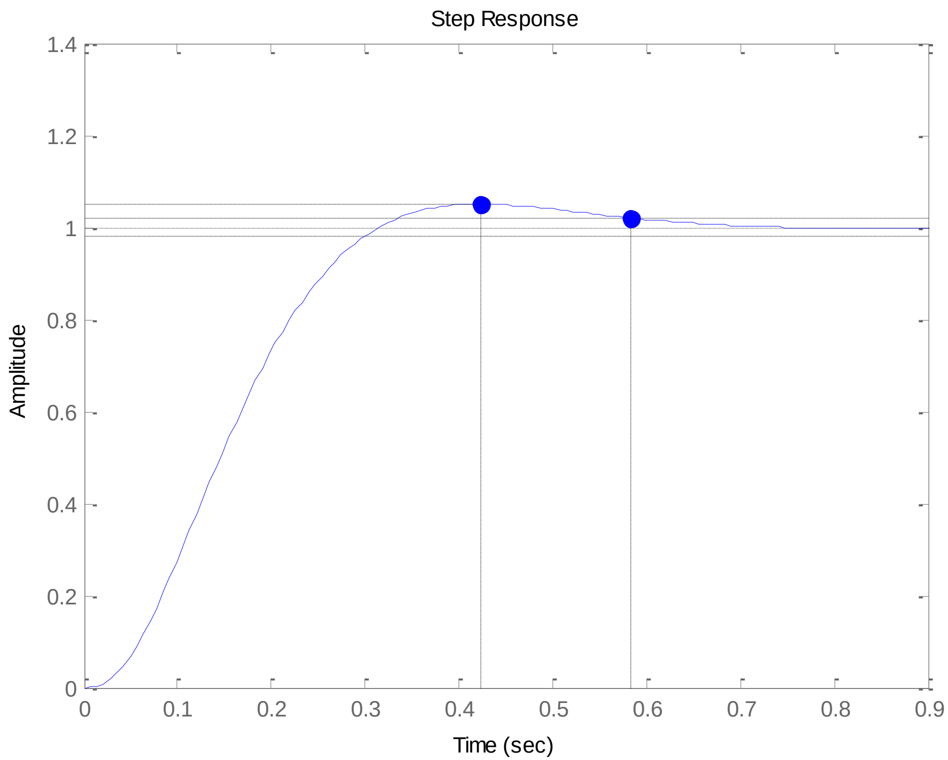
Example 2.3



Of course the response just not jump to unity in zero time, there is a transient. The transient is comprised of on rapidly decaying exponentially which also has a low weighting (this is essentially negligible), one somewhat rapidly decaying exponential which is a little more heavily weighted (this affects the transient response, particularly in the early phase) and one oscillatory term having a frequency of 7.3661 rad/sec and having a somewhat rapidly exponentially decaying amplitude. The latter term is the dominant transient term.

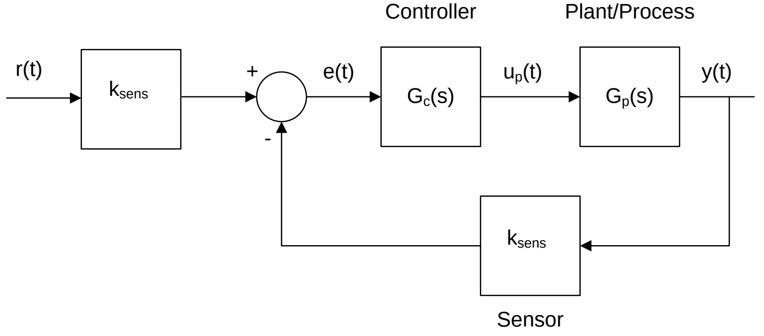
Example 2.3

The Matlab command **step** can be used to calculate and plot the step response.



We confirm a slightly oscillatory transient.

Example 2.3



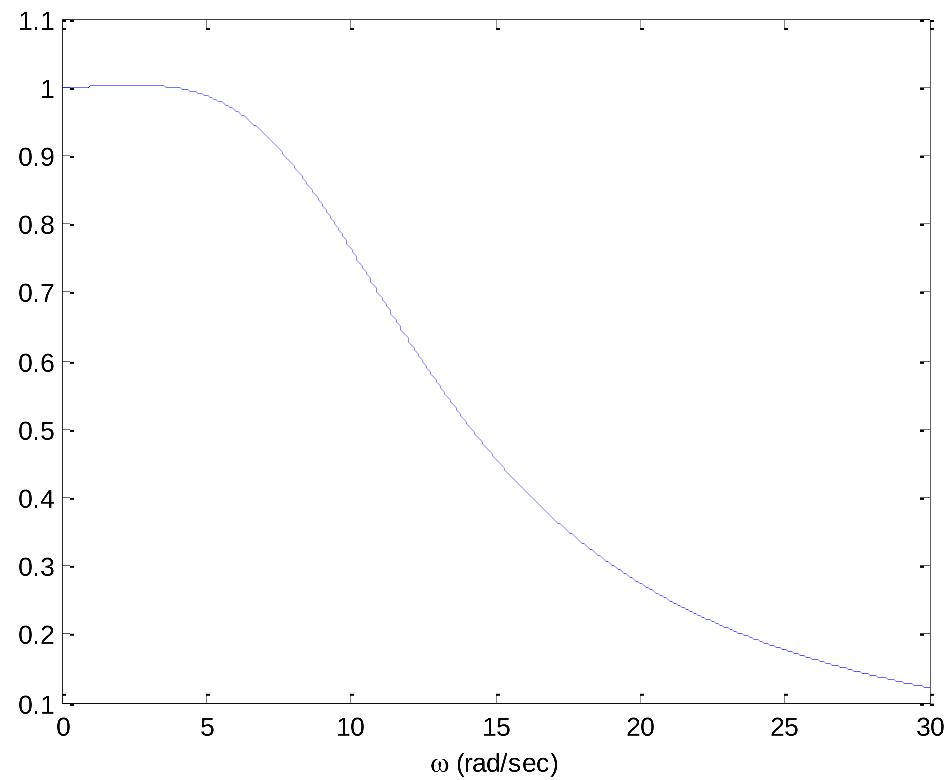
The second very important response is the response to a co-sinusoid. We have seen that this depends upon the transfer function evaluated at $s = j\omega$ where ω is the frequency of the sinusoid, i.e. on the complex number:

$$\frac{8{,}000(j\omega)+88{,}000}{\left((j\omega)^4+92(j\omega)^3+2{,}060(j\omega)^2+20{,}400(j\omega)+88{,}000\right)}$$

Example 2.3

The magnitude of this number determines the gain which a co-sinusoid experiences. The argument determines the phase-shift. We may plot the magnitude response.

The plot reveals a slight resonance at a frequency of about 2.67 rad/sec. The system passes low frequencies and blocks high frequencies.

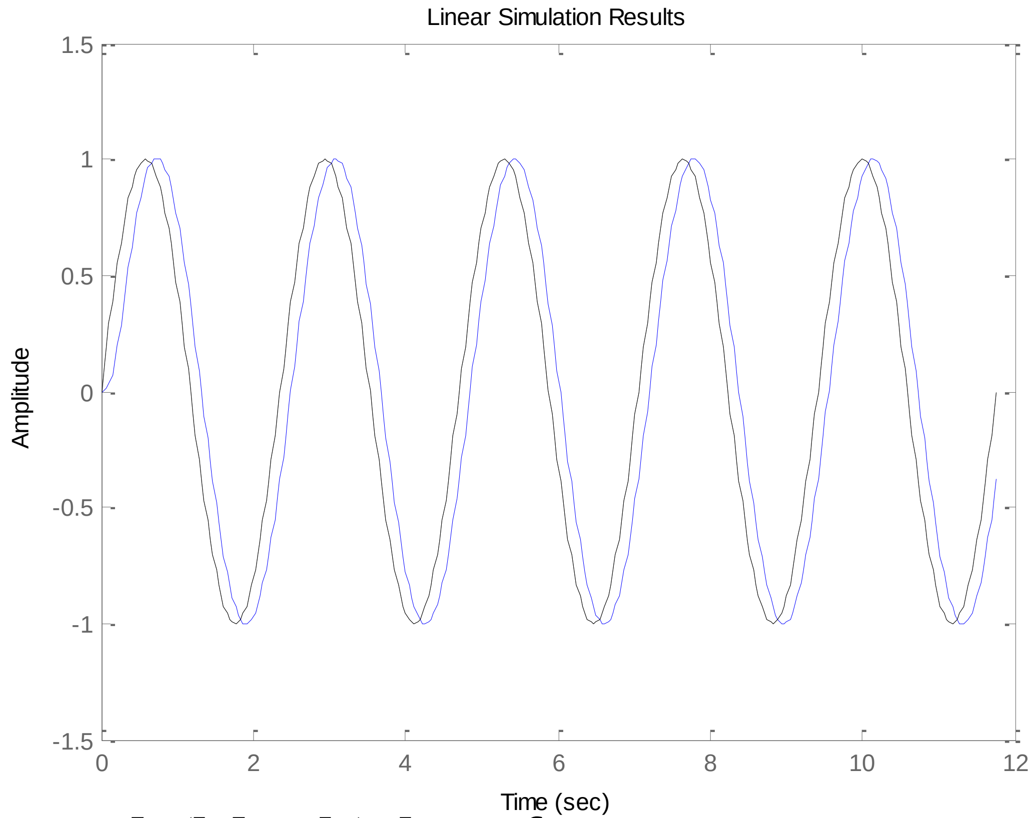


Example 2.3

```
>>  Num = [8000 88000]          create numerator polynomial Num

>>  Den = [1 92 2060 20400 88000]    create denominator polynomial Den

>> w = [0:0.01:30];      create vector of frequencies of interest

>> GF = polyval(Num,i*w)./polyval(Den,i*w)   evaluate transfer function at s = jw for all of these frequencies

>> plot(w,abs(GF))   plot magnitude response

>>  xlabel('\omega (rad/sec)')   label horizontal axis
```

Example 2.3

Matlab confirms this result:



When the input signal (black) has frequency 2.67 rad/sec the output (blue) amplitude is virtually identical and the phase shift is small.

Example 2.3

>> Num = [8000 88000] *create numerator polynomial Num*

>> Den = [1 92 2060 20400 88000] *create denominator polynomial Den*

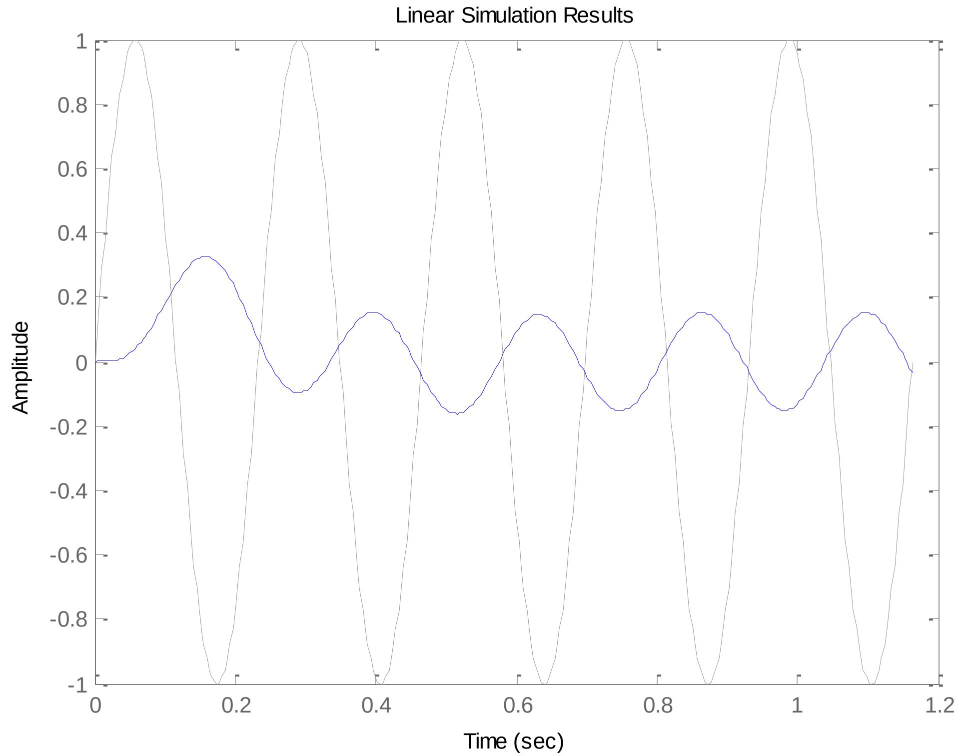
>> G = tf(Num,Den) *create transfer function/system variable*

>> [U,T] = gensig('sin',2*pi/2.67); *create time and amplitude data for unit amplitude sinuosoid of period $2\pi/2.67$*

>> lsim(G,U,T) *plot response to sinusoidal input. Output is plotted in blue and input in black*

Example 2.3

Matlab confirms this result:



When the input signal (black) has frequency 27 rad/sec the output (blue) amplitude is much smaller and the phase shift is close to 180 degrees.