Signals and Systems

Robby McKilliam

August 27, 2013

Contents

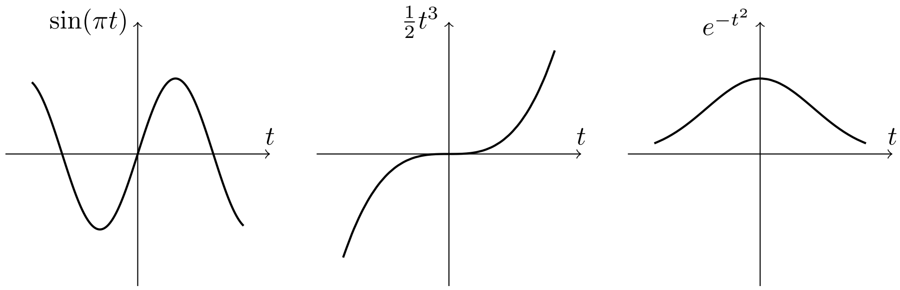

Figure 1: 1-dimensional continuous-time signals

1 Signals and systems

A **signal** is a function mapping an input variable to some output variable. For example

$$\sin(\pi t), \qquad \frac{1}{2}t^3, \qquad e^{-t^2}$$

all represent **signals** with input variable $t \in \mathbb{R}$, and they are plotted in Figure 1. If x is a signal and t an input variable we write $x(t)$ for the output variable. Signals can be multidimensional. This page is an example of a 2-dimensional signal, the independent variables are the horizontal and vertical position on the page, and the signal maps this position to a colour, in this case either black or white. A moving image such as seen on your television or computer monitor is an example of a 3-dimensional signal, the three independent variables being vertical and horizontal screen position and time. The signal maps each position and time to a colour on the screen. In this course we focus exclusively on 1-dimensional signals such as those in Figure 1 and we will only consider signals that are real or complex valued. Many of the results presented here can be extended to deal with multidimensional signals.

1.1 Properties of signals

A signal x is **bounded** if there exists a real number M such that

$$|x(t)| \leq M \qquad \text{for all } t \in \mathbb{R}$$

where $|\cdot|$ denotes the (complex) magnitude. Both $\sin(\pi t)$ and e^{-t^2} are examples of bounded signals because $|\sin(\pi t)| \leq 1$ and $|e^{-t^2}| \leq 1$ for all $t \in \mathbb{R}$. However, $\frac{1}{2}t^3$ is not bounded because its magnitude grows indefinitely as t moves away from the origin.

A signal x is **periodic** if there exists a real number T such that

$$x(t) = x(t + kT) \qquad \text{for all } k \in \mathbb{Z} \text{ and } t \in \mathbb{R}.$$

For example, the signal $\sin(\pi t)$ is periodic with period $T = 2$. Neither $\frac{1}{2}t^3$ or e^{-t^2} are periodic.

A signal x is called **locally integrable** if for all constants a and b,

$$\int_a^b |x(t)|\, dt$$

exists (evaluates to a finite number). An example of a signal that is not locally integrable is $x(t) = \frac{1}{t}$ (Exercise 1.2). Two signals x and y are equal, i.e. $x = y$ if $x(t) = y(t)$ for all $t \in \mathbb{R}$.

A signal x is called **absolutely integrable** if

$$\|x\|_1 = \int_{-\infty}^{\infty} |x(t)|\, dt \tag{1.1}$$

exists. Here we introduce the notation $\|x\|_1$ called the ℓ_1**-norm** of x. For example $\sin(\pi t)$ and $\frac{1}{2}t^3$ are not absolutely integrable, but e^{-t^2} is because [Nicholas and Yates, 1950]

$$\int_{-\infty}^{\infty} |e^{-t^2}| dt = \int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\pi}. \tag{1.2}$$

The signal x is called is **square integrable** if

$$\|x\|_2 = \int_{-\infty}^{\infty} |x(t)|^2\, dt$$

exists. Square integrable signals are also called **energy signals**, and the value of $\|x\|_2$ is called the **energy** of x (it is also called the ℓ_2**-norm** of x). For example $\sin(\pi t)$ and $\frac{1}{2}t^3$ are not energy signals, but e^{-t^2} is.

A signal x is **right sided** if there exists a $T \in \mathbb{R}$ such that $x(t) = 0$ for all $t < T$. Correspondingly x is **left sided** if $x(t) = 0$ for all $T > t$. For example, the **step function**

$$u(t) = \begin{cases} 1 & t > 0, \\ 0 & t \leq 0 \end{cases} \tag{1.3}$$

is right-sided. Its reflection in time $u(-t)$ is left sided (Figure 2). A signal x is called **finite in time** if it is both left and right sided, that is, if there exits a $T \in \mathbb{R}$ such that $x(t) = x(-t) = 0$ for all $t > T$. A signal is called **unbounded in time** if it is neither left nor right sided. For example, the continuous time signals $\sin(\pi t)$ and e^{-t^2} are unbounded in time, but the **rectangular pulse**

$$\Pi(t) = \begin{cases} 1 & -\frac{1}{2} < t \leq \frac{1}{2} \\ 0 & \text{otherwise}, \end{cases} \tag{1.4}$$

is finite in time.

1.2 Systems (functions of signals)

A **system** (also known as an **operator** or **functional**) maps a signal to another signal. For example

$$x(t) + 3x(t-1), \qquad \int_0^1 x(t-\tau) d\tau, \qquad \frac{1}{x(t)}, \qquad \frac{d}{dt}x(t)$$

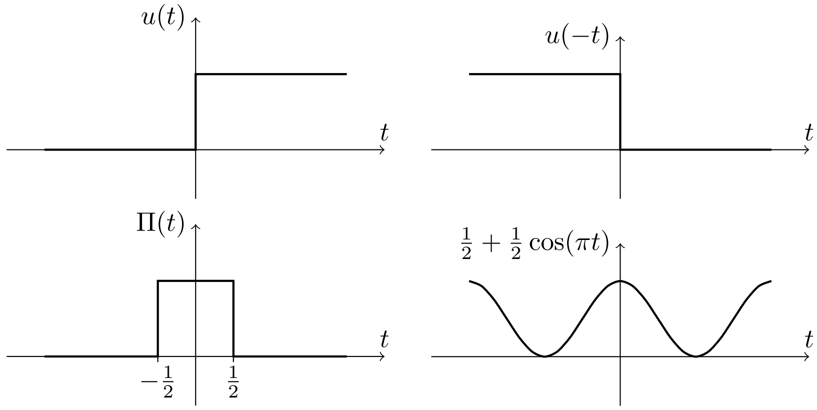

Figure 2: The right sided step function $u(t)$, its left sided reflection $u(-t)$, the finite in time rectangular pulse $\Pi(t)$ and the unbounded in time signal $\frac{1}{2} + \frac{1}{2}\cos(x)$.

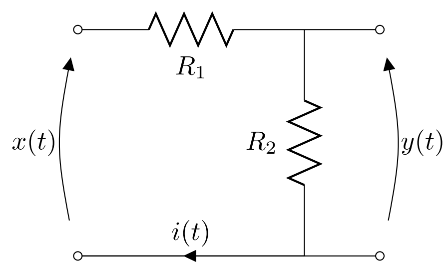

Figure 3: A **voltage divider** circuit.

represent systems, each mapping the signal x to another signal. Consider the electric circuit in Figure 3 called a **voltage divider**. If the voltage at time t is $x(t)$ then, by Ohm's law, the current at time t satisfies

$$i(t) = \frac{1}{R_1 + R_2} x(t),$$

and the voltage over the resistor R_2 is

$$y(t) = R_2 i(t) = \frac{R_2}{R_1 + R_2} x(t) \tag{1.5}$$

The circuit can be considered as a system mapping the signal x representing the voltage to the signal $i = \frac{1}{R_1+R_2}x$ representing the current, or a system mapping x to the signal $y = \frac{R_2}{R_1+R_2}x$ representing the voltage over resistor R_2.

We denote systems with capital letters such as H and G. A system H is a function that maps a signal x to another signal denoted $H(x)$. We call x the **input signal** and $H(x)$ the **output signal** or the **response** of system H to signal x. If we want to include the independent variable t we will write $H(x)(t)$

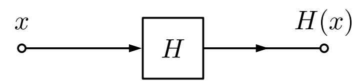

Figure 4: System block diagram with input signal x and output signal $H(x)$.

or $H(x, t)$ and do not distinguish between these [Curry and Feys, 1968]. It is sometimes useful to depict systems with a block diagram. Figure 4 is a simple block diagram showing the input and output signals of a system H.

Using this notation the electric circuit in Figure 3 corresponds with the system

$$H(x) = \frac{R_2}{R_1 + R_2} x = y.$$

This system multiplies the input signal x by $\frac{R_2}{R_1+R_2}$. This brings us to our first practical test.

Test 1 (**Voltage divider**) In this test we construct the voltage divider from Figure 3 on a breadboard with resistors $R_1 \approx 100\Omega$ and $R_2 \approx 470\Omega$ with values accurate to within 5%. Using a computer soundcard (an approximation of) the voltage signal

$$x(t) = \sin(2\pi f_1 t) \qquad \text{with} \qquad f_1 = 100$$

is passed through the circuit. The approximation is generated by sampling $x(t)$ at rate $F_s = \frac{1}{T_s} = 44100\text{Hz}$ to generate samples

$$x_n = x(nT_s) \qquad n = 0, \dots, 2F_s$$

corresponding to approximately 2 seconds of signal. These samples are passed to the soundcard which starts playback. The voltage over the resistor R_2 is recorded (also using the soundcard) that returns a lists of samples $y_1, \dots, y_L$ taken at rate F_s. The continuous-time voltage over R_2 can be (approximately) reconstructed from these samples as

$$\tilde{y}(t) = \sum_{\ell=1}^{L} y_\ell \operatorname{sinc}(F_s t - \ell) \tag{1.6}$$

where

$$\operatorname{sinc}(t) = \frac{\sin(\pi t)}{\pi t} \tag{1.7}$$

is the called the **sinc function** and is plotted in Figure 6. We will justify this reconstruction in Section 6. Simultaneously the (stereo) soundcard is used to record the input voltage $x(t)$ producing samples $x_1, \dots, x_L$ taken at rate F_s. An approximation of the continuous-time input signal is

$$\tilde{x}(t) = \sum_{\ell=1}^{L} x_\ell \operatorname{sinc}(F_s t - \ell). \tag{1.8}$$

In view of (1.5) we would expect the approximate relationship

$$\tilde{y} \approx \frac{R_2}{R_1 + R_2}\tilde{x} = \frac{42}{57}\tilde{x}$$

A plot of $\tilde{y}$, $\tilde{x}$ and $\frac{42}{57}\tilde{x}$ over a 20ms period from 1s to 1.02s is given in Figure 5. The hypothesised output signal $\frac{42}{57}\tilde{x}$ does not match the observed output signal $\tilde{y}$. A primary reason is that the circuitry inside the soundcard itself cannot be ignored. When deriving the equation for the voltage divider we implicitly assumed that current flows through the output of the soundcard without resistance (a short circuit), and that no current flows through the input device of the soundcard (an open circuit). These assumptions are not realistic. Modelling the circuitry in the sound card wont be attempted here. In the next section we will construct circuits that contain external sources of power (active circuits). These are less sensitive to the circuitry inside the soundcard.

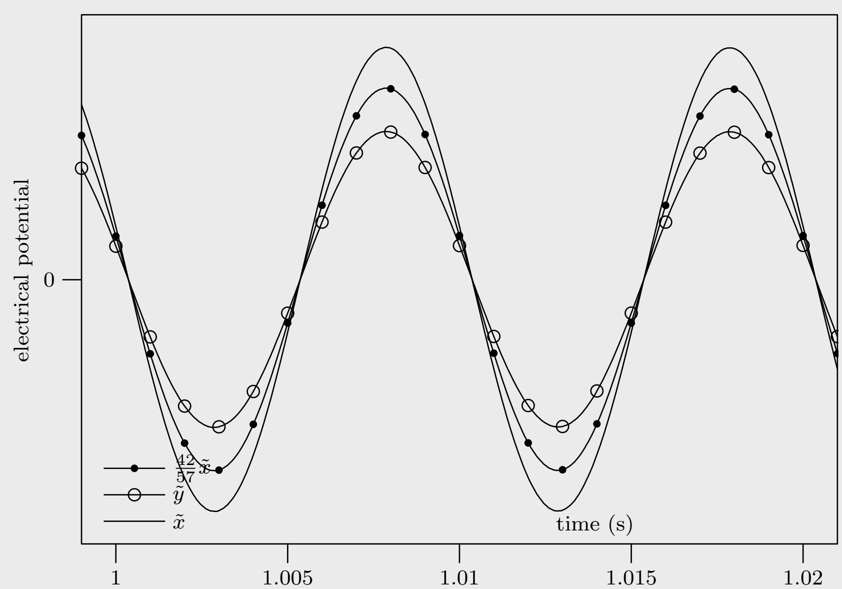

Figure 5: Plot of reconstructed input signal $\tilde{x}$ (solid line), output signal $\tilde{y}$ (solid line with circle) and hypothesised output signal $\frac{42}{57}\tilde{x}$ (solid line with dot) for the voltage divider circuit in Figure 3. The hypothesised signal does not match $\tilde{y}$. One reason is that the model does not take account of the circuitry inside the soundcard.

Not all signals can be input to all systems. For example, the system

$$H(x, t) = \frac{1}{x(t)}$$

is not defined at those t where $x(t) = 0$ because we cannot divide by zero.

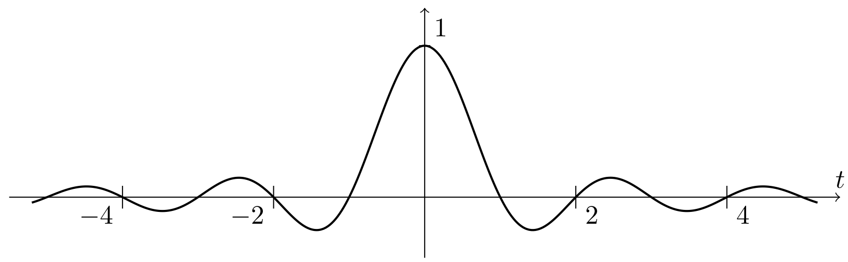

Figure 6: The **sinc function** $\mathrm{sinc}(t) = \frac{\sin(\pi t)}{\pi t}$.

Another example is the system

$$I_\infty(x, t) = \int_{-\infty}^{t} x(\tau) d\tau, \tag{1.9}$$

called an **integrator**, that is not defined for those signals where the integral above does not exist (is not finite). For example, the signal $x(t) = 1$ cannot be input to the integrator since the integral $\int_{-\infty}^{t} dt$ does not exist.

Thus, when specifying a system it is necessary to also specify a set of signals that can be input, called the **domain** of the system. For example, the domain of the system $H(x, t) = \frac{1}{x(t)}$ is the set of signals $x(t)$ which are not zero for any t. The domain of the integrator $I_\infty(x, t)$ is the set of signals for which the integral $\int_{-\infty}^{t} x(\tau) d\tau$ exists for all $t \in \mathbb{R}$. The domain of a system is usually obvious from the specification of the system itself. For this reason we will not usually state the domain explicitly. We will only do so if there is chance for confusion.

1.3 Some important systems

The system

$$T_\tau(x, t) = x(t - \tau)$$

is called the **time-shifter**. This system shifts the input signal along the t axis ('time' axis) by τ. When τ is positive T_τ delays the input signal by τ. The time-shifter will appear so regularly in this course that we use the special notation T_τ to represent it. Figure 7 depicts the action of time-shifters $T_{1.5}$ and T_{-3} on the signal $x(t) = e^{-t^2}$. When $\tau = 0$ the time-shifter is the **identity system**

$$T_0(x) = x$$

that maps the signal x to itself.

Another important system is the **time-scaler** that has the form

$$H(x, t) = x(\alpha t)$$

for $\alpha \in \mathbb{R}$. Figure 8 depicts the action of a time-scaler with a number of values for α. When $\alpha = -1$ the time-scaler reflects the input signal in the time axis.

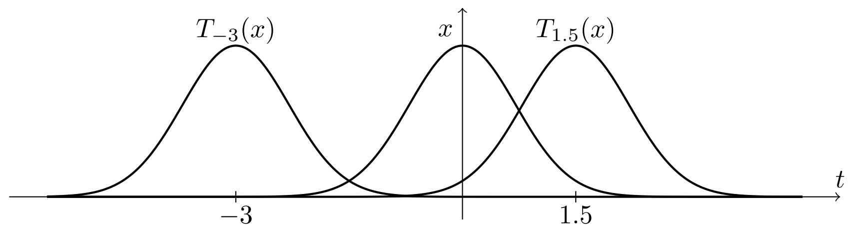

Figure 7: Time-shifter system $T_{1.5}(x,t) = x(t-1.5)$ and $T_{-3}(x,t) = x(t+3)$ acting on the signal $x(t) = e^{-t^2}$.

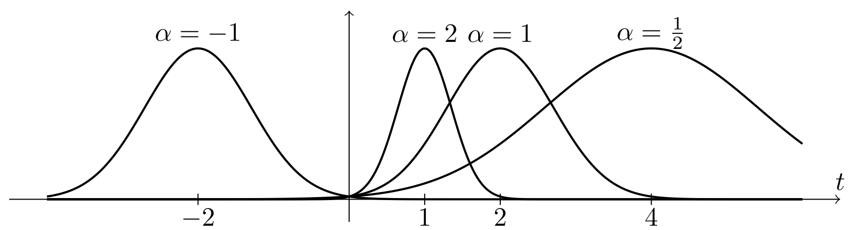

Figure 8: Time-scaler system $H(x,t) = x(\alpha t)$ for $\alpha = -1, \frac{1}{2}, 1$ and 2 acting on the signal $x(t) = e^{-(t-2)^2}$.

Another system we regularly encounter is the **differentiator**

$$D(x,t) = \frac{d}{dt}x(t),$$

that returns the derivative of the input signal. We also define a kth differentiator

$$D^k(x,t) = \frac{d^k}{dt^k}x(t)$$

that returns the kth derivative of the input signal.

Another important system is the **integrator**

$$I_a(x,t) = \int_{-a}^{t} x(\tau)d\tau.$$

The parameter a describes the lower bound of the integral. In this course it will often be that $a = \infty$ or $a = 0$. The integrator can only be applied to those signals for which the integral above exists. For example, the integrator I_∞ can be applied to the signal $tu(t)$ where $u(t)$ is the step function (1.3). The output signal is

$$\int_{-\infty}^{t} \tau u(\tau)d\tau = \begin{cases} \int_0^t \tau d\tau = \frac{t^2}{2} & t > 0 \\ 0 & t \leq 0. \end{cases}$$

However, the integrator cannot be applied to the signal $x(t) = t$ because $\int_{-\infty}^{t} \tau d\tau$ does not exist.

1.4 Properties of systems

A system H is called **memoryless** if the output signal $H(x)$ at time t depends only on the input signal x at time t. For example $\frac{1}{x(t)}$ and the identity system T_0 are memoryless, but

$$x(t) + 3x(t-1) \qquad \text{and} \qquad \int_0^1 x(t-\tau)d\tau$$

are not. A time-shifter system T_τ with $\tau \neq 0$ is not memoryless.

A system H is **causal** if the output signal $H(x)$ at time t depends on the input signal only at times less than or equal to t. Memoryless systems such as $\frac{1}{x(t)}$ and T_0 are also causal. Time-shifters $T_\tau(x,t) = x(t-\tau)$ are causal when $\tau \geq 0$, but are not causal when $\tau < 0$. The systems

$$x(t) + 3x(t-1) \qquad \text{and} \qquad \int_0^1 x(t-\tau)d\tau$$

are causal, but the systems

$$x(t) + 3x(t+1) \qquad \text{and} \qquad \int_0^1 x(t+\tau)d\tau$$

are not causal.

A system H is called **bounded-input-bounded-output (BIBO) stable** or just **stable** if the output signal $H(x)$ is bounded whenever the input signal x is bounded. That is, H is stable if for every positive real number M there exists a positive real number K such that for all signals x satisfying

$$|x(t)| < M \qquad \text{for all } t \in \mathbb{R},$$

it also holds that

$$|H(x,t)| < K \qquad \text{for all } t \in \mathbb{R}.$$

For example, the system $x(t) + 3x(t-1)$ is stable with $K = 4M$ since if $|x(t)| < M$ then

$$|x(t) + 3x(t-1)| \leq |x(t)| + 3\,|x(t-1)| < 4M = K.$$

The integrator I_a for any $a \in \mathbb{R}$ and differentiator D are not stable (Exercises 1.5 and 1.6).

A system H is **linear** if

$$H(ax + by) = aH(x) + bH(y)$$

for all signals x and y, and for all complex numbers a and b. That is, a linear system has the property: If the input consists of a weighted sum of signals, then the output consists of the same weighted sum of the responses of the system to

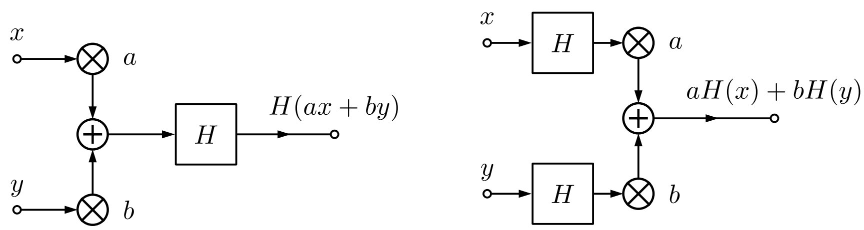

Figure 9: If H is a linear system the outputs of these two diagrams are the same signal, i.e. $H(ax+by) = aH(x) + bH(y)$.

those signals. Figure 9 indicates the linearity property using a block diagram. For example, the differentiator is linear because

$$\begin{aligned} D(ax+by,t) &= \frac{d}{dt}\big(ax(t)+by(t)\big) \\ &= a\frac{d}{dt}x(t) + b\frac{d}{dt}y(t) \\ &= aD(x,t) + bD(y,t), \end{aligned}$$

but the system $H(x,t) = \frac{1}{x(t)}$ is not linear because

$$H(ax+by,t) = \frac{1}{ax(t)+by(t)} \neq \frac{a}{x(t)} + \frac{b}{y(t)} = aH(x,t) + bH(y,t)$$

in general.

The property of linearity trivially generalises to more than two signals. For example if $x_1, \dots, x_k$ are signals and $a_1, \dots, a_k$ are complex numbers for some finite k, then

$$H(a_1x_1 + \dots + a_kx_k) = a_1H(x_1) + \dots + a_kH(x_k).$$

A system H is **time-invariant** if

$$H\big(T_\tau(x), t\big) = H(x, t-\tau)$$

for all signals x and all time-shifts $\tau \in \mathbb{R}$. That is, a system is time-invariant if time-shifting the input signal results in the same time-shift of the output signal. Equivalently, H is time-invariant if H commutes the time-shifter T_τ, that is, if

$$H\big(T_\tau(x)\big) = T_\tau\big(H(x)\big)$$

for all $\tau \in \mathbb{R}$ and all signals x. Figure 10 represents the property of time-invariance with a block diagram.

Let S be a set of signals. A system H is said to be **invertible** on S if each signal $x \in S$ is mapped to a unique signal $H(x)$. That is, for all signals $x, y \in S$ then $H(x) = H(y)$ if and only if $x = y$. If a system H is invertible on S then there exists an inverse system H^{-1} such that

$$x = H^{-1}\big(H(x)\big) \qquad \text{for all } x \in S.$$

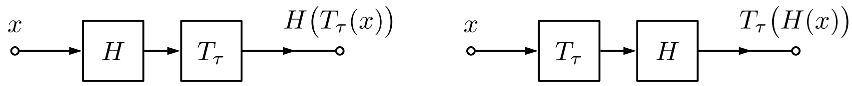

Figure 10: If H is a time-invariant system the outputs of these two diagrams are the same signal, i.e. $H\big(T_\tau(x)\big) = T_\tau\big(H(x)\big)$.

For example, let S be a any set of signals. The time-shifter T_τ is invertible on S. The inverse system is $T_{-\tau}$ since

$$T_{-\tau}\big(T_\tau(x), t\big) = x(t - \tau + \tau) = x(t).$$

As another example, let S be the set of differentiable signals. The differentiator system D is **not** invertible on S because if $x \in S$ and if $y(t) = x(t) + c$ for any constant c then $D(y) = D(x)$. However, if we restrict S to those differentiable signals for which $x(0) = c$ is fixed, then D *is* invertible on S. The inverse system in this case is

$$D^{-1}(x, t) = I_0(x, t) + c = \int_0^t x(t)dt + c$$

because

$$D^{-1}(D(x), t) = \int_0^t D(x, t)dt + c = \int_0^t \frac{d}{dt}x(t)dt + x(0) = x(t)$$

by the fundamental theorem of calculus.

1.5 Exercises

1.1. State whether the step function $u(t)$ is bounded, periodic, absolutely summable, an energy signal.

1.2. Show that the signal t^2 is locally integrable, but that the signal $\frac{1}{t^2}$ is not.

1.3. Plot the signal

$$x(t) = \begin{cases} \frac{1}{t+1} & t > 0 \\ \frac{1}{t-1} & t \leq 0. \end{cases}$$

State whether it is: bounded, locally integrable, absolutely integrable, square integrable.

1.4. Compute the energy of the signal $e^{-\alpha^2 t^2}$ (Hint: use equation (1.2) on page 4 and a change of variables).

1.5. Show that the integrator I_a for any $a \in \mathbb{R}$ is not stable.

1.6. Show that the differentiator system D is not stable.

1.7. Show that the time-shifter is linear and time-invariant, and that the time-scaler is linear, but not time invariant

1.8. Show that the integrator I_c with c finite is linear, but not time-invariant.

1.9. Show that the integrator I_∞ is linear and time invariant.

1.10. State whether the system $H(x,t) = x(t) + 1$ is linear, time-invariant, stable.

1.11. State whether the system $H(x,t) = 0$ is linear, time-invariant, stable.

1.12. State whether the system $H(x,t) = 1$ is linear, time-invariant, stable.

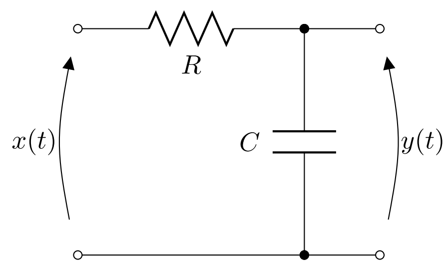

Figure 11: An electrical circuit with resistor and capacitor in series, otherwise known as an **RC circuit**.

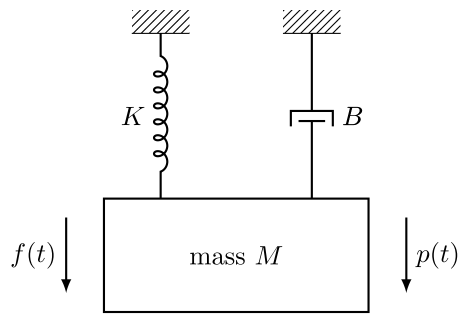

Figure 12: A mechanical mass-spring-damper system

2 Systems modelled by differential equations

Systems of significant interest in this course are those where the input signal x and output signal y are related by a linear differential equation with constant coefficients, that is, an equation of the form

$$\sum_{\ell=0}^{m} a_\ell \frac{d^\ell}{dt^\ell} x(t) = \sum_{\ell=0}^{k} b_\ell \frac{d^\ell}{dt^\ell} y(t)$$

where $a_0, \dots, a_m$ and $b_0, \dots, b_k$ are constant real numbers. In what follows we will use the differentiator system $D(x)$ rather than the notation $\frac{d}{d\ell} x(t)$ to represent differentiation of the signal x. To represent the ℓth derivative we write $D^\ell(x)$. Using this notation the differential equation above is

$$\sum_{\ell=0}^{m} a_\ell D^\ell(x) = \sum_{\ell=0}^{k} b_\ell D^\ell(y). \tag{2.1}$$

Equations of this form can be used to model a large number of electrical, mechanical and other real world devices. For example, consider the resistor and capacitor (RC) circuit in Figure 11. Let the signal v_R represent the voltage over the resistor and i the current through both resistor and capacitor. The voltage signals satisfy

$$x = y + v_R,$$

and the current satisfies both

$$v_R = Ri, \qquad \text{and} \qquad i = CD(y).$$

Combining these equations,

$$x = y + RCD(y) \tag{2.2}$$

that is in the form of (2.1).

As another example, consider the mass, spring and damper in Figure 12. A force represented by the signal f is externally applied to the mass, and the position of the mass is represented by the signal p. The spring exerts force $-Kp$ that is proportional to the position of the mass, and the damper exerts force $-BD(p)$ that is proportional to the velocity of the mass. The cumulative force exerted on the mass is

$$f_m = f - Kp - BD(p)$$

and by Newton's law the acceleration of the mass $D^2(p)$ satisfies

$$MD^2(p) = f_m = f - Kp - BD(p),$$

from which we obtain the differential equation

$$f = Kp + BD(p) + MD^2(p) \tag{2.3}$$

that is in the form of (2.1) if we put $x = f$ and $y = p$. Given p we can readily solve for the corresponding force f. As a concrete example, let the spring constant, damping constant and mass be $K = B = M = 1$. If the position satisfies $p(t) = e^{-t^2}$, then the corresponding force satisfies

$$f(t) = e^{-t^2}(4t^2 - 2t - 1).$$

Figure 13 depicts these signals.

What happens if a particular force signal f is applied to the mass? For example, say we apply the force

$$f(t) = \Pi(t - \tfrac{1}{2}) = \begin{cases} 1 & 0 < t \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

What is the corresponding position signal p? We are not yet ready to answer this question, but will be later (Exercise 4.8).

In both the mechanical mass-spring-damper system in Figure 12 and the electrical RC circuit in Figure 11 we obtain a differential equation relating the input signal x with the output signal y. The equations do not specify the output signal y explicitly in terms of the input signal x, that is, they do not explicitly define a system H such $y = H(x)$. As they are, the differential equations,

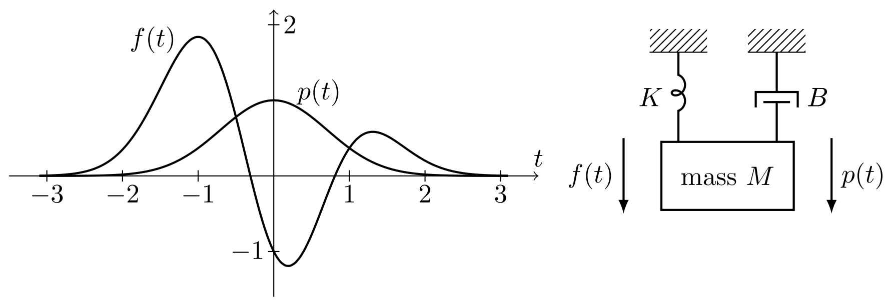

Figure 13: A solution to the mass spring damper system with $K = B = M = 1$. The position is $p(t) = e^{-t^2}$ with corresponding force $f(t) = e^{-t^2}(4t^2 - 2t - 1)$.

do not provide as much information about the behaviour of the system as we would like. For example, is the system stable? Is it invertible? The **Laplace transform**, described in Section 4, is a useful tool for answering these questions. A key property enabling the Laplace transform is that differential equations of the form (2.1) describe systems that are linear and time-invariant.We further study linear, time-invariant systems in Section 3. The remainder of this section details the construction of differential equations that model various mechanical, electrical, and electro-mechanical systems. We will use the systems constructed here as examples throughout the course.

2.1 Passive circuits

Passive electrical circuits require no sources of power other than the input signal itself. For example, the voltage divider in Figure 3 and the RC circuit in Figure 11 are passive circuits. Another common passive electrical circuit is the resistor, capacitor and inductor (RLC) circuit depicted in Figure 14. In this circuit we let the output signal y be the voltage over the resistor. Let v_C represent the voltage over the capacitor and v_L the voltage over the inductor and let i be the current. We have

$$y = Ri, \qquad i = CD(v_C), \qquad v_L = LD(i),$$

leading to the following relationships between y, v_C and v_L,

$$y = RCD(v_C), \qquad Rv_L = LD(y).$$

Kirchhoff's voltage law gives $x = y + v_C + v_L$ and by differentiating both sides

$$D(x) = D(y) + D(v_C) + D(v_L).$$

Substituting the equations relating y, v_C and v_L leads to

$$RCD(x) = y + RCD(y) + LCD^2(y). \tag{2.4}$$

We can similarly find equations relating the input voltage with v_C and v_L.

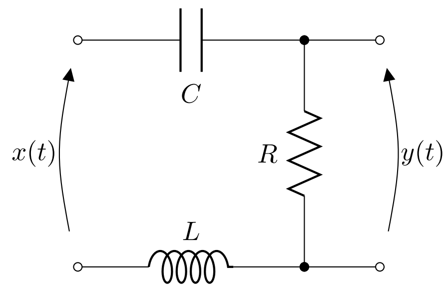

Figure 14: An electrical circuit with resistor, capacitor and inductor in series, otherwise known as an **RLC circuit**.

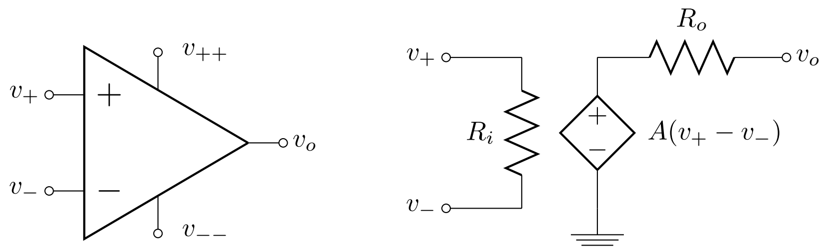

Figure 15: Left: triangular component diagram of an **operational amplifier**. The v_{++} and v_{--} connectors indicate where an external voltage source can be connected to the amplifier. These connectors will usually be omitted. Right: model for an operational amplifier including input resistance R_i, output resistance R_o, and open loop gain A. The diamond shaped component is a dependent voltage source. This model is only useful when the operational amplifier is in a negative feedback circuit.

2.2 Active circuits

Unlike passive electrical circuits, an **active circuit** requires a source of power external to the input signal. In this course active circuits will be modelled and constructed using **operational amplifiers** as depicted in Figure 15. The left hand side of Figure 15 shows a triangular circuit diagram for an operational amplifier, and the right hand side of Figure 15 shows a circuit that can be used to model the behaviour of the amplifier. The v_{++} and v_{--} connectors indicate where an external voltage source can be connected to the amplifier, and will normally not be drawn. The diamond shaped component is a dependent voltage source with voltage $A(v_+ - v_-)$ that depends on the difference between the voltage at the **non-inverting input** v_+ and the voltage at the **inverting input** v_-. The dimensionless constant A is called the **open loop gain**. Most operational amplifiers have large open loop gain A, large input resistance R_i and small output resistance R_o. As we will see, it can be convenient to consider the behaviour as $A \to \infty$, $R_i \to \infty$ and $R_0 \to 0$, resulting in an **ideal operational amplifier**.

As an example, an operational amplifier configured as a **multiplier** is depicted in Figure 16. This circuit is an example of an operation amplifier configured with **negative feedback**, meaning that the output of the amplifier is connected (in this case by a resistor) to the inverting input. The horizontal wire at the bottom of the plot is consider to be ground (zero volts) and is connected to the negative terminal of the dependent voltage source of the operational amplifier depicted in Figure 15. An equivalent circuit for the multiplier using the model in Figure 15 is shown in Figure 17. Solving this circuit (Exercise 2.1) yields the following relationship between the input voltage signal x and the output voltage signal y,

$$y = \frac{R_i(AR_2 + R_o)}{R_i(R_2 + R_o) + R_1(R_2 + R_i - AR_i + R_o)} x. \tag{2.5}$$

For an ideal operational amplifier we let $A \to \infty$, $R_i \to \infty$ and $R_o \to 0$. In this case terms involving the product AR_i dominate and we are left with the simpler equation

$$y = -\frac{R_2}{R_1} x. \tag{2.6}$$

Thus, assuming an ideal operational amplifier, the circuit acts as a multiplier with constant $-\frac{R_2}{R_1}$.

The equation relating x and y is much simpler for the ideal operational amplifier. Fortunately this equation can be obtained directly using the following two rules:

1. the voltage at the inverting and non-inverting inputs are equal,

2. no current flows through the inverting and non-inverting inputs.

These rules are only useful for analysing circuits with negative feedback. Let us now rederive (2.6) using these rules. Since the non-inverting input is connected to ground, the voltage at the inverting input is zero. So, the voltage over resistor R_2 is $y = R_2 i$. Since no current flows through the inverting input the current through R_1 is also i and $x = -R_1 i$. Combing these results, the input voltage x and the output voltage y are related by

$$y = -\frac{R_2}{R_1} x.$$

In Test 2 the inverting amplifier circuit is constructed and the relationship above is tested using a computer soundcard.

We now consider another circuit consisting of an operational amplifier, two resistors and a capacitor depicted in Figure 18. Assuming an ideal operational amplifier, the voltage at the inverting terminal is zero because the non-inverting terminal is connected to ground. Thus, the voltage over capacitor C_2 and resistor R_2 is equal to y and, by Kirchoff's current law

$$i = \frac{y}{R_2} + C_2 D(y).$$

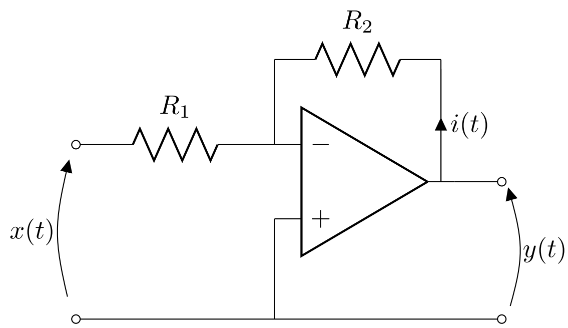

Figure 16: Inverting amplifier

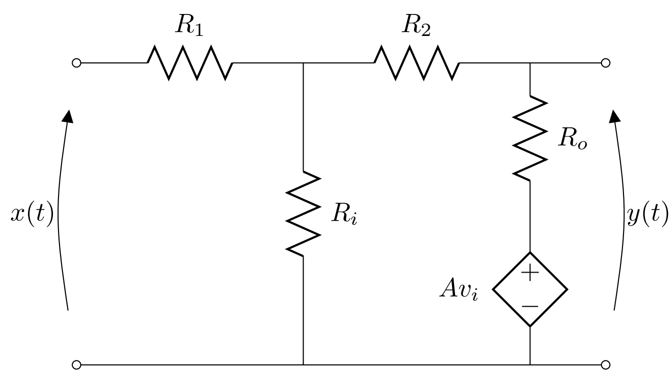

Figure 17: An equivalent circuit for the inverting amplifier from Figure 16 using the model for an operational amplifier in Figure 15. The symbol v_i is the voltage over resistor R_i.

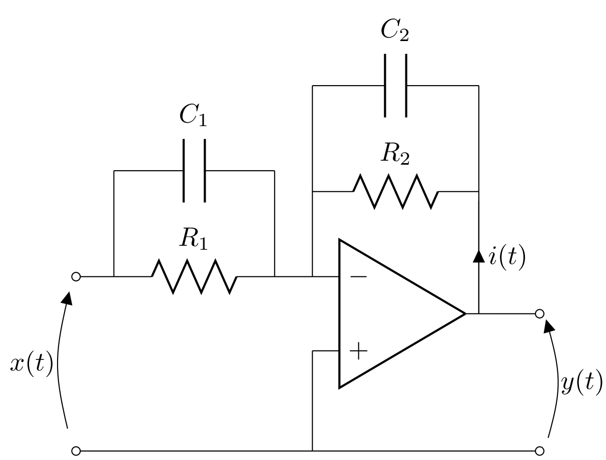

Figure 18: Operational amplifier configured with two capacitors and two resistors.

Test 2 (**Inverting amplifier**) In this test we construct the inverting amplifier circuit from Figure 16 with $R_2 \approx 22\text{k}\Omega$ and $R_1 \approx 12\text{k}\Omega$ that are accurate to within 5% of these values. The operational amplifier used is the Texas Instruments LM358P. Using a computer soundcard (an approximation of) the voltage signal

$$x(t) = \tfrac{1}{3}\sin(2\pi f_1 t) + \tfrac{1}{3}\sin(2\pi f_2 t)$$

with $f_1 = 100$ and $f_2 = 233$ is passed through the circuit. As in previous tests, the soundcard is used to sample the input signal x and the output signal y. Approximate reconstructions of the input signal $\tilde{x}$ and output signal $\tilde{y}$ are given according to (1.8), and (1.6). According to (2.4) we expect the approximate relationship

$$\tilde{y} \approx -\frac{R_2}{R_1}\tilde{x} = -\tfrac{11}{6}\tilde{x}.$$

Each of $\tilde{y}$, $\tilde{x}$ and $-\frac{11}{6}\tilde{x}$ are plotted in Figure 19. Observe that the amplitude of the hypothesised output signal $-\frac{11}{6}\tilde{x}$ is slightly larger than the observed output signal $\tilde{y}$. One explanation is that the ideal model we have used for the operational amplifier is only an approximation.

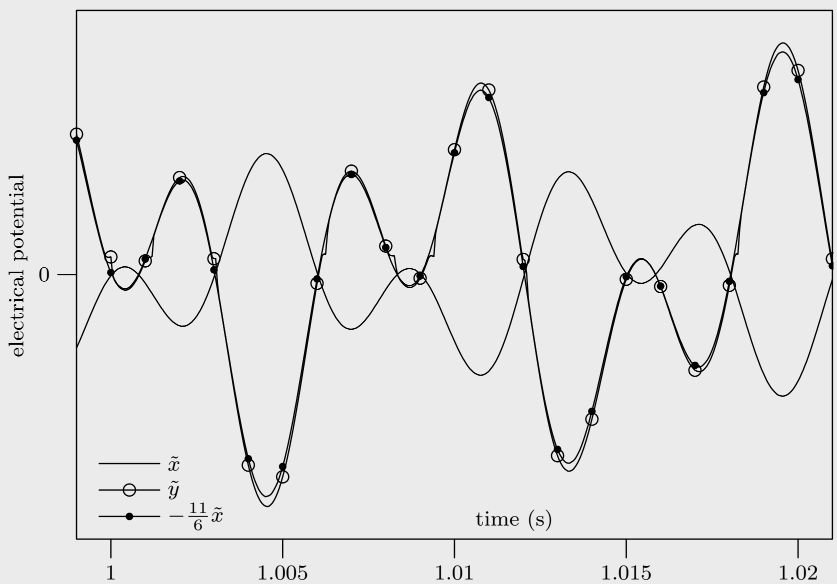

Figure 19: Plot of reconstructed input signal $\tilde{x}$ (solid line), output signal $\tilde{y}$ (solid line with circle) and hypothesised output signal $-\frac{11}{6}\tilde{x}$ (solid line with dot).

Similarly, since no current flows through the inverting terminal,

$$i = -\frac{x}{R_1} - C_1 D(x).$$

Combining these equations yields

$$-\frac{x}{R_1} - C_1 D(x) = \frac{y}{R_2} + CD(y). \tag{2.7}$$

Observe the similarity between this equation and that for the passive RC circuit (2.2) when $R_1 = R_2$ and $C_1 = 0$ (an open circuit). In this case

$$x = -y - R_1 C_2 D(y). \tag{2.8}$$

This circuit is tested in Test 3.

Test 3 (**Active RC circuit**) In this test we construct the circuit from Figure 18 with $R_1 \approx R_2 \approx 27\text{k}\Omega$ and $C_2 \approx 10\text{nF}$ accurate to within 5% of these values and $C_1 = 0$ (an open circuit). The operational amplifier used is a Texas Instruments LM358P. Using a computer soundcard (an approximation of) the voltage signal

$$x(t) = \tfrac{1}{3}\sin(2\pi f_1 t) + \tfrac{1}{3}\sin(2\pi f_2 t)$$

with $f_1 = 500$ and $f_2 = 1333$ is passed through the circuit. As in previous tests, the soundcard is used to sample the input signal x and the output signal y and approximate reconstructions $\tilde{x}$ and $\tilde{y}$ are given according to (1.8) and (1.6). According to (2.8) we expect the approximate relationship

$$\tilde{x} \approx -\frac{R_1}{R_2}\tilde{y} - R_1 C D(\tilde{y}) = -\tilde{y} - \tfrac{27}{10000} D(\tilde{y}).$$

The derivative of the sinc function is

$$D(\text{sinc}, t) = \frac{1}{\pi t^2}\big(\pi t \cos(\pi t) - \sin(\pi t)\big), \tag{2.9}$$

and so,

$$D(\tilde{y}) = D\left(\sum_{\ell=1}^{L} y_\ell \,\text{sinc}(F_s t - \ell)\right) = F_s \sum_{\ell=1}^{L} y_\ell D(\text{sinc}, F_s t - \ell). \tag{2.10}$$

Each of $\tilde{y}$, $\tilde{x}$ and $-\tilde{y} - \tfrac{27}{10000} D(\tilde{y})$ are plotted in Figure 19.

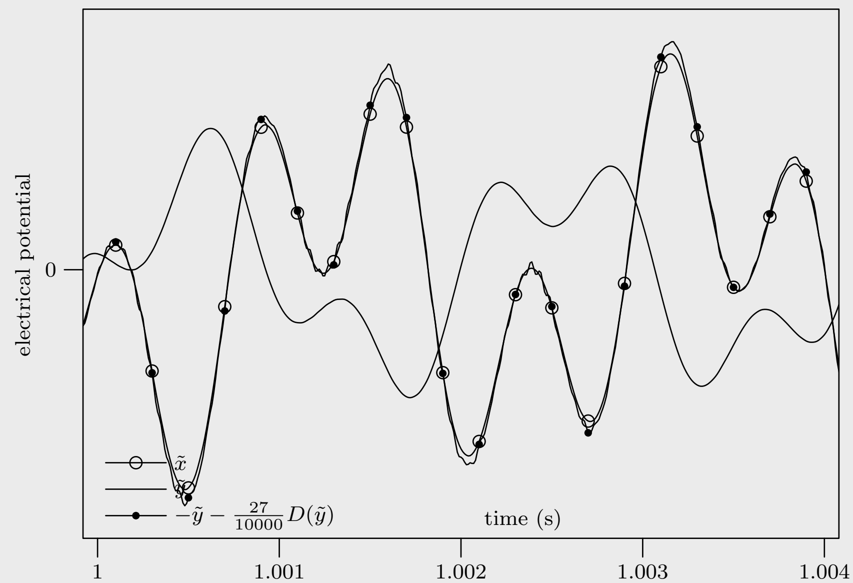

Figure 19: Plot of reconstructed input signal $\tilde{x}$ (solid line with circle), output signal $\tilde{y}$ (solid line), and hypothesised input signal $-\tilde{y} - \frac{27}{10000}D(\tilde{y})$ (solid line with dot).

Consider the circuit in Figure 20. Assuming an ideal operational amplifier, the input voltage x satisfies

$$-i = \frac{x}{R_1} + C_1 D(x).$$

The voltage over the capacitor C_2 is $y - R_2 i$ and so the current satisfies

$$i = C_2 D(y - R_2 i).$$

Combining these equations gives

$$-\frac{x}{R_1} - C_1 D(x) = C_2 D(y) + \frac{R_2 C_2}{R_1} D(x) + R_2 C_2 C_1 D^2(x),$$

and after rearranging,

$$D(y) = -\frac{1}{R_1 C_1} x - \left(\frac{R_2}{R_1} + \frac{C_1}{C_2} \right) D(x) - R_2 C_1 D^2(x).$$

Put

$$K_i = \frac{1}{R_1 C_2}, \qquad K_p = \frac{R_2}{R_1} + \frac{C_1}{C_2}, \qquad K_d = R_2 C_1$$

and now

$$D(y) = -K_i x - K_p D(x) - K_d D^2(x). \tag{2.11}$$

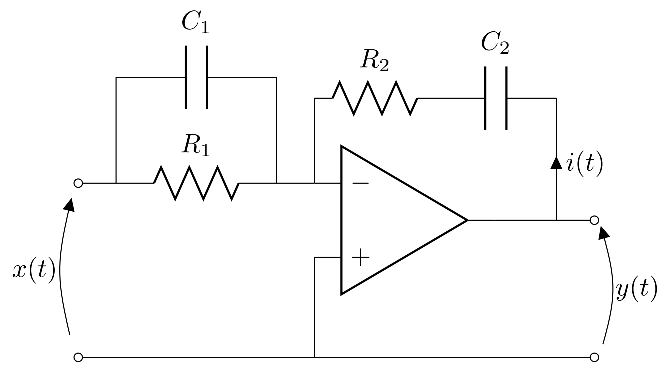

Figure 20: Operational amplifier implementing a **proportional-integral-derivative controller**.

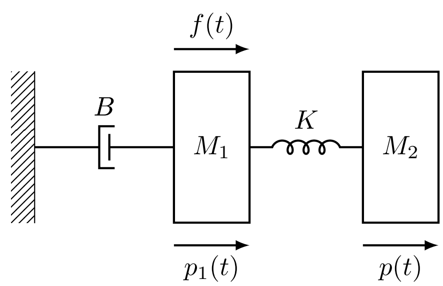

Figure 21: Two masses, a spring and a damper

This equation models what is called a **proportional-integral-derivative controller** or **PID controller**. The coefficients K_i, K_p and K_d are called the **integral gain**, **proportional gain**, and **derivative gain**.

2.3 Masses, springs and dampers

A mechanical mass, spring, damper system was described in Section 2 and Figure 12 We now consider another mechanical system involving a different configuration of masses, a spring and a damper depicted in Figure 21. A mass M_1 is connected to a wall by a damper with constant B, and to another mass M_2 by a spring with constant K. A force represented by the signal f is applied to the first mass. We will derive a differential equation relating f with the position p of the second mass. We assume that the spring applies no force (is in equilibrium) when masses are distance d apart. The forces due to the spring satisfy

$$f_{s1} = -f_{s2} = K(p - p_1 - d)$$

where f_{s1} and f_{s2} are signals representing the force due to the spring on mass M_1 and M_2 respectively. It is convenient to define the signal $g(t) = p_1(t) + d$ so that forces due to spring satisfy the simpler equation

$$f_{s1} = -f_{s2} = K(p - g).$$

The only force applied to M_2 is by the spring and so, by Newton's law, the acceleration of M_2 satisfies

$$M_2 D^2(p) = f_{s2}.$$

Substituting this into the previous equation gives a differential equation relating g and p,

$$Kg = Kp + M_2 D^2(p). \tag{2.12}$$

The force applied by the damper on mass M_1 is given by the signal

$$f_d = -BD(p_1) = -BD(g)$$

where the replacement of p_1 by g is justified because differentiation will remove the constant d. The cumulative force on M_1 is given by the signal

$$\begin{aligned} f_1 &= f + f_d + f_{s1} \\ &= f - Kg + Kp - BD(g), \end{aligned} \tag{2.13}$$

and by Newton's law the acceleration of M_1 satisfies

$$M_1 D^2(p_1) = M_1 D^2(g) = f_1.$$

Substituting this into (2.13) and using (2.12) we obtain a fourth order differential equation relating p and f,

$$f = BD(p) + (M_1 + M_2)D^2(p) - \frac{BM_2}{K} D^3(p) + \frac{M_1 M_2}{K} D^4(p). \tag{2.14}$$

Given the position of the second mass p we can readily solve for the corresponding force f and position of the first mass p. For example, if the constants $B = K = 1$ and $M_1 = M_2 = \frac{1}{2}$ and $d = \frac{5}{2}$, and if the position of the second mass satisfies

$$p(t) = e^{-t^2}$$

then, by application of (2.14) and (2.12),

$$f(t) = e^{-t^2}(1 - 8t - 8t^2 + 4t^3 + 4t^4), \qquad \text{and} \qquad p_1(t) = 2e^{-t^2}t^2 - \tfrac{5}{2}.$$

This solution is plotted in Figure 22.

2.4 Direct current motors

Direct current (DC) motors convert electrical energy, in the form of a voltage, into rotary kinetic energy [Nise, 2007, page 76]. We derive a differential equation relating the input voltage v to the angular position of the motor θ. Figure 23 depicts the components of a DC motor.

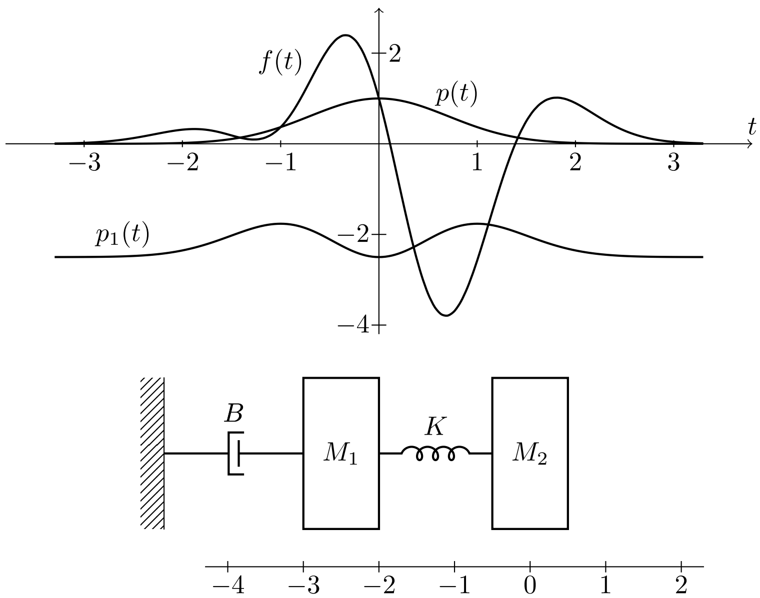

Figure 22: Solution of the system describing two masses with a spring and damper where $B = K = 1$ and $M_1 = M_2 = \frac{1}{2}$ and the position of the second mass is $p(t) = e^{-t^2}$.

The voltages over the resistor and inductor satisfy

$$v_R = Ri, \qquad v_L = LD(i),$$

and the motion of the motor induces a voltage called the back electromotive force (EMF),

$$v_b = K_b D(\theta)$$

that we model as being proportional to the angular velocity of the motor. The input voltage now satisfies

$$v = v_R + v_L + v_b = Ri + LD(i) + K_b D(\theta).$$

The torque τ applied by the motor is modelled as being proportional to the current i,

$$\tau = K_\tau i.$$

A load with inertia J is attached to the motor. Two forces are assumed to act on the load, the torque τ applied by the current, and a torque $\tau_d = BD(\theta)$ modelling a damper that acts proportionally against the angular velocity of the motor. By Newton's law, the angular acceleration of the load satisfies

$$JD^2(\theta) = \tau - \tau_d = K_\tau i - BD(\theta).$$

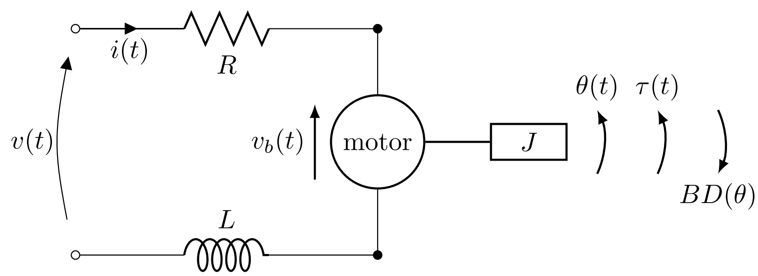

Figure 23: Diagram for a rotary direct current (DC) motor

Combining these equations we obtain the 3rd order differential equation

$$v = \left(\frac{RB}{K_\tau} + K_b\right) D(\theta) + \frac{RJ + LB}{K_\tau} D^2(\theta) + \frac{LJ}{K_\tau} D^3(\theta)$$

relating voltage and motor position. In many DC motors the inductance L is small and can be ignored, leaving the simpler second order equation

$$v = \left(\frac{RB}{K_\tau} + K_b\right) D(\theta) + \frac{RJ}{K_\tau} D^2(\theta). \tag{2.15}$$

Given the position signal θ we can find the corresponding voltage signal v. For example, put the constants $K_b = K_\tau = B = R = J = 1$ and assume that

$$\theta(t) = 2\pi(1 + \operatorname{erf}(t))$$

where $\operatorname{erf}(t) = \frac{2}{\pi}\int_{-\infty}^{t} e^{-\tau^2} d\tau$ is the **error function**. The corresponding angular velocity $D(\theta)$ and voltage v satisfy

$$D(\theta, t) = 4\sqrt{\pi}e^{-t^2}, \qquad v(t) = 8\sqrt{\pi}e^{-t^2}(1 - t).$$

These signals are depicted in Figure 24. This voltage signal is sufficient to make the motor perform two revolutions and then come to rest.

2.5 Exercises

2.1. Analyse the inverting amplifier circuit in Figure 17 to obtain the relationship between input voltage x and output voltage y given by (2.5). You may wish to use a symbolic programming language, for example Mathematica.

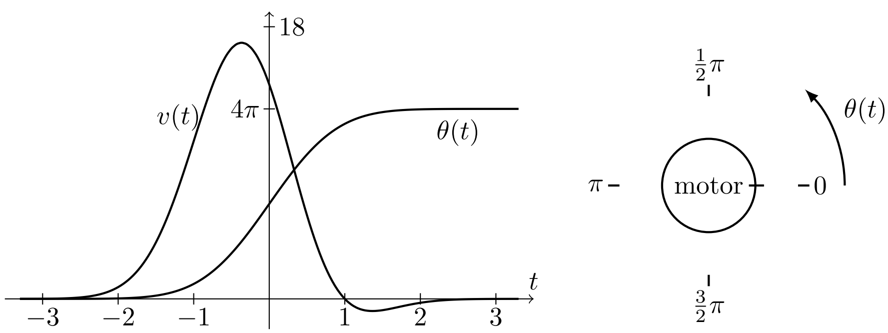

Figure 24: Voltage and corresponding angle for a DC motor with constants $K_b = K_\tau = B = R = J = 1$.

3 Linear time-invariant systems

Throughout this section we let H be a linear time-invariant system.

3.1 Convolution, regular systems and the delta "function"

A large number of linear time-invariant systems can be represented by a signal called the **impulse response**. The impulse response of a system H is a signal h such that

$$H(x,t) = \int_{-\infty}^{\infty} h(\tau)x(t-\tau)d\tau,$$

that is, the response of H to input signal x can be represented as an integral equation involving x and the impulse response h. The integral is called a **convolution** and appears so often a special notation is used for it

$$h * x = \int_{-\infty}^{\infty} h(\tau)x(t-\tau)d\tau.$$

Those systems that have an impulse response we call **regular systems**[1]. Observe that regular systems are linear because

$$\begin{aligned} H(ax+by) &= h*(ax+by) \\ &= \int_{-\infty}^{\infty} h(\tau)\big(ax(t-\tau)+by(t-\tau)\big)d\tau \\ &= a\int_{-\infty}^{\infty} h(\tau)x(t-\tau)d\tau + b\int_{-\infty}^{\infty} h(\tau)y(t-\tau)d\tau \\ &= a(h*x)+b(h*y) \\ &= aH(x)+bH(y). \end{aligned} \tag{3.1}$$

The above equations show that convolution commutes with scalar multiplication and distributes with addition, that is

$$h*(ax+by) = a(h*x)+b(h*y).$$

Regular systems are also time-invariant because

$$\begin{aligned} T_\kappa\big(H(x)\big) &= H(x,t-\kappa) \\ &= \int_{-\infty}^{\infty} h(\tau)x(t-\kappa-\tau)d\tau \\ &= \int_{-\infty}^{\infty} h(\tau)T_\kappa(x,t-\tau)d\tau \\ &= H\big(T_\kappa(x)\big). \end{aligned}$$

[1] The name **regular system** is motivated by the term **regular distribution** [Zemanian, 1965]

We can define the impulse response of a regular system H in the following way. First define the signal

$$p_\gamma(t) = \begin{cases} \gamma, & 0 < t \leq \frac{1}{\gamma} \\ 0, & \text{otherwise} \end{cases}$$

that is a rectangular shaped pulse of height γ and width $\frac{1}{\gamma}$. The signal p_γ is plotted in Figure 25 for $\gamma = \frac{1}{2}, 1, 2, 5$. As γ increases the pulse gets thinner and higher so as to keep the area under p_γ equal to one. Consider the response of the regular system H to the signal p_γ,

$$\begin{aligned} H(p_\gamma) &= h * p_\gamma \\ &= \int_{-\infty}^{\infty} h(\tau) p_\gamma(t - \tau) d\tau \\ &= \gamma \int_0^{1/\gamma} h(\tau) d\tau. \end{aligned}$$

Taking limits as $\gamma \to \infty$,

$$\lim_{\gamma \to \infty} H(p_\gamma) = \lim_{\gamma \to \infty} \gamma \int_0^{1/\gamma} h(\tau) d\tau = h.$$

Thus, we define the impulse response of a regular system is H as the limit $h = \lim_{\gamma \to \infty} H(p_\gamma)$. The limit exists when H is regular. If this limit does not exist, the system is not regular and does not have an impulse response.

As an example, consider the integrator system

$$I_\infty(x, t) = \int_{-\infty}^{t} x(\tau) d\tau \tag{3.2}$$

described in Section 1.3. This systems response to p_γ is

$$I_\infty(p_\gamma, t) = \int_{-\infty}^{t} p_\gamma(\tau) d\tau = \begin{cases} 0, & t \leq 0 \\ \gamma t, & 0 < t \leq \frac{1}{\gamma} \\ 1, & t > \frac{1}{\gamma}. \end{cases}$$

The response is plotted in Figure 25. Taking the limit as $\gamma \to \infty$ we find that the impulse response of the integrator is the step function

$$u(t) = \lim_{\gamma \to \infty} H(p_\gamma) = \begin{cases} 0 & t \leq 0 \\ 1 & t > 0. \end{cases} \tag{3.3}$$

Some important systems do not have an impulse response. For example, the identity system T_0 does not because

$$\lim_{\gamma \to \infty} T_0(p_\gamma) = \lim_{\gamma \to \infty} p_\gamma$$

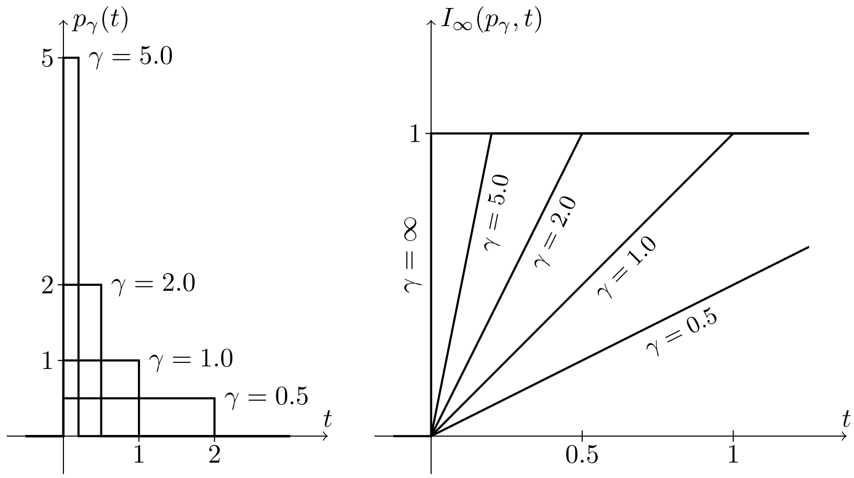

Figure 25: The rectangular shaped pulse p_γ for $\gamma = 0.5, 1, 2, 5$ and the response of the integrator (3.2) to p_γ for $\gamma = 0.5, 1, 2, 5, \infty$.

does not exist. Similarly, all the time shifters T_τ do not have impulse responses. However, it is notationally useful to pretend that T_0 *does* have an impulse response and we denote it by the symbol δ called the **delta function**. The idea is to assign δ the property

$$\int_{\infty}^{\infty} x(t)\delta(t)dt = x(0)$$

so that convolution of x and δ is

$$\delta * x = \int_{-\infty}^{\infty} \delta(\tau)x(t-\tau)d\tau = x(t) = T_0(x,t).$$

We now treat δ as if it were a signal. So $\delta(t-\tau)$ will represent the impulse response of the time shifter T_τ because

$$\begin{aligned} T_\tau(x) &= \delta(t-\tau) * x \\ &= \int_{-\infty}^{\infty} \delta(\kappa-\tau)x(t-\kappa)d\kappa \\ &= \int_{-\infty}^{\infty} \delta(k)x(t-\tau-k)dk \qquad\qquad \text{(change variable } k = \kappa-\tau) \\ &= x(t-\tau). \end{aligned}$$

It is important to realise that δ is not actually a signal. It is not a function. However, it can be convenient to treat δ as if it were a function. The manipulations in the last set of equations, such as the change of variables, are not formally justified, but they do lead to the desired result $T_\tau(x) = x(t-\tau)$ in this case. There is no guarantee that mechanical mathematical manipulations involving δ will lead to sensible results in general.

The only other non regular systems that we have use of are differentiators D^k, and it is convenient to define a similar notation for pretending that these systems have an impulse response. In this case we use the symbol δ^k and assign it to have the property

$$\int_{\infty}^{\infty} x(t)\delta^k(t)dt = D^k(x, 0),$$

so that convolution of x and δ is

$$\delta^k * x = \int_{-\infty}^{\infty} \delta^k(\tau)x(t-\tau)d\tau = D^k(x, t).$$

As with the delta function the symbol δ^k must be treated with care. This notation can be useful, but purely formal manipulations with δ^k may not lead to sensible results in general.

The impulse response h immediately yields some properties of the corresponding system H. For example, if $h(t) = 0$ for all $t < 0$, then H is causal because

$$H(x, t) = h * x = \int_{-\infty}^{\infty} h(\tau)x(t-\tau)d\tau = \int_{0}^{\infty} h(\tau)x(t-\tau)d\tau$$

only depends on values of x at times less than t, i.e., only times $t-\tau$ with $\tau > 0$. The system H is stable if and only if h is absolutely integrable (Exercise 3.3).

Another important signal is the **step response** of a system that is defined as the response of the system to the step function $u(t)$. For example, the step response of the time shifter T_τ is the time shifted step function $T_\tau(u, t) = u(t-\tau)$. The step response of the integrator I_∞ is

$$I_\infty(u, t) = \int_{-\infty}^{t} u(t)dt = \begin{cases} \int_0^t dt = t & t > 0 \\ 0 & t \leq 0. \end{cases}$$

This signal is often called the **ramp function**. Not all systems have a step response. For example, the system with impulse response $u(-t)$ does not because the convolution of the step $u(t)$ and it's reflection $u(-t)$ does not exist. If a system H has both an impulse respone h and a step response $H(u)$, then these two signals are related. To see this, observe that the step response is

$$H(u) = h * u = \int_{-\infty}^{\infty} h(\tau)u(t-\tau)d\tau = \int_{-\infty}^{t} h(\tau)d\tau = I_\infty(h, t). \tag{3.4}$$

Thus, the step response can be obtained by applying the integrator I_∞ to the impulse response.

3.2 Properties of convolution

The convolution $x*y$ of two signals x and y does not always exist. For example, if $x = u(t)$ and $y = u(-t)$, then

$$x * y = \int_{-\infty}^{\infty} u(\tau)u(\tau - t)d\tau = \int_{t}^{\infty} d\tau,$$

which is not finite for any t. On the other hand, if $x = y = u(t)$, then

$$x * y = \int_{-\infty}^{\infty} u(\tau)u(t - \tau)d\tau = \begin{cases} \int_0^t dt = t & t > 0 \\ 0 & t \leq 0, \end{cases}$$

which exists for all t.

We have already shown in (3.1) that convolution commutes with scalar multiplication and is distributive with addition, that is, for signals x, y, w and complex numbers a, b,

$$a(x * w) + b(y * w) = (ax + by) * w.$$

Convolution is commutative, that is, $x*y = y*x$ whenever these convolutions exist. To see this, write

$$\begin{aligned} x * y &= \int_{-\infty}^{\infty} x(\tau)y(t - \tau)d\tau \\ &= \int_{-\infty}^{\infty} x(t - \kappa)y(\kappa)d\kappa \qquad \text{(change variable } \kappa = t - \tau\text{)} \\ &= y * x. \end{aligned}$$

Convolution is also associative, that is, for signals x, y, z,

$$(x * y) * z = x * (y * z). \qquad \text{(see Exercise 3.2)}$$

By combining the associative and commutative properties we find that the order in which the convolutions in $x * y * z$ are performed does not mater, that is

$$x * y * z = y * z * x = z * x * y = y * x * z = x * z * y = z * y * x$$

provided that all the convolutions involved exist. More generally, the order in which any sequence of convolutions is performed does not change the final result.

3.3 Linear combining and composition

Let H_1 and H_2 be linear time-invariant systems and let H be the system

$$H(x) = cH_1(x) + dH_2(x), \qquad c, d \in \mathbb{R}$$

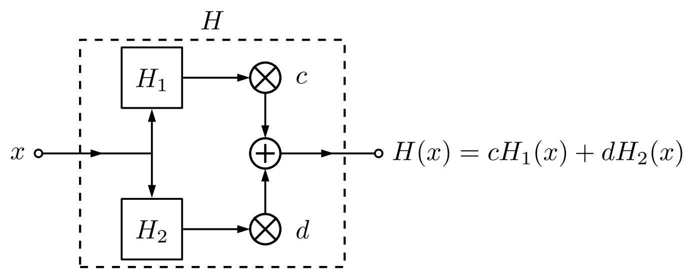

Figure 26: Block diagram depicting the linear combining property of linear time-invariant systems. The system $cH_1(x) + dH_2(x)$ can be expressed as a single linear time-invariant system $H(x)$.

formed by a linear combination of H_1 and H_2. The system H is linear because for signals x, y and complex numbers a, b,

$$\begin{aligned}
H(ax+by) &= cH_1(ax+by) + dH_2(ax+by) \\
&= acH_1(x) + bcH_1(y) + adH_2(x) + bdH_2(y) \qquad \text{(linearity } H_1, H_2\text{)} \\
&= a\big(cH_1(x) + dH_2(x)\big) + b\big(cH_1(y) + dH_2(y)\big) \\
&= aH(x) + bH(y).
\end{aligned}$$

The system is also time-invariant because

$$\begin{aligned}
H\big(T_\tau(x)\big) &= cH_1\big(T_\tau(x)\big) + dH_2\big(T_\tau(x)\big) \\
&= cT_\tau\big(H_1(x)\big) + dT_\tau\big(H_2(x)\big) \qquad \text{(time-invariance } H_1, H_2\text{)} \\
&= T_\tau\big(cH_1(x) + dH_2(x)\big) \qquad \text{(linearity } T_\tau\text{)} \\
&= T_\tau\big(H(x)\big).
\end{aligned}$$

So, we can construct linear time-invariant systems by **linearly combining** (adding and multiplying by constants) other linear time-invariant systems. If H_1 and H_2 are regular systems this linear combining property can be expressed using their impulse responses h_1 and h_2. We have

$$\begin{aligned}
H(x) &= aH_1(x) + bH_2(x) \\
&= ah_1 * x + bh_2 * x \\
&= (ah_1 + bh_2) * x \qquad \text{(distributivity of convolution)} \\
&= h * x,
\end{aligned}$$

and so, H is a regular system with impulse response $h = ah_1 + bh_2$.

Another way to construct linear time-invariant systems is by **composition**. Let H_1 and H_2 be linear time-invariant systems and let

$$H(x) = H_2\big(H_1(x)\big),$$

that is, H first applies the system H_1 and then applies the system H_2. The composition $H_2\big(H_1(x)\big)$ only applies to those signals x in the domain of H_1 and

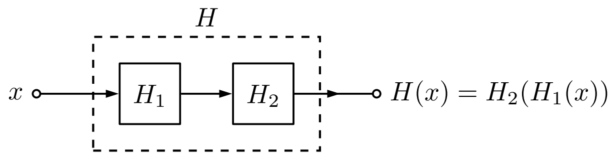

Figure 27: Block diagram depicting the composition property of linear time-invariant systems. The system $H_2(H_1(x))$ can be expressed as a single linear time-invariant system $H(x)$.

such that the signal $H_1(x)$ is in the domain of H_2. The system H is linear because, for signals x, y and complex numbers a, b,

$$\begin{aligned}
H(ax+by) &= H_2\big(H_1(ax+by)\big) \\
&= H_2\big(aH_1(x)+bH_1(y)\big) && \text{(linearity } H_1\text{)} \\
&= aH_2\big(H_1(x)\big)+bH_2\big(H_1(y)\big)\big) && \text{(linearity } H_2\text{)} \\
&= aH(x)+bH(y).
\end{aligned}$$

The system is also time-invariant because

$$\begin{aligned}
H\big(T_\tau(x)\big) &= H_2\big(H_1(T_\tau(x)\big) \\
&= H_2\big(T_\tau\big(H_1(x)\big)\big) && \text{(time-invariance } H_1\text{)} \\
&= T_\tau\big(H_2\big(H_1(x)\big)\big) && \text{(time-invariance } H_2\text{)} \\
&= T_\tau\big(H(x)\big).
\end{aligned}$$

If H_1 and H_2 are regular systems the composition property can be expressed using their impulse responses h_1 and h_2. It follows that

$$\begin{aligned}
H(x) &= H_2(H_1(x)) \\
&= h_2 * (h_1 * x) \\
&= (h_2 * h_1) * x && \text{(associativity of convolution)} \\
&= h * x,
\end{aligned}$$

and so, H is a regular system with impulse response $h = h_2 * h_1$. Thus, if H is the composition of regular systems H_1 and H_2, then H is a regular system with impulse given by the convolution of the impulse responses of H_1 and H_2.

A wide variety of linear time-invariant systems by now be contructed by linearly combining and composing simpler systems.

3.4 Eigenfunctions and the transfer function

Let $s = \sigma + j\omega \in \mathbb{C}$. Complex exponential signals of the form

$$e^{st} = e^{\sigma t} e^{j\omega t} = e^{\sigma t}\big(\cos(\omega t) + j\sin(\omega t)\big)$$

play an important role in the study of linear time-invariant systems. Let H be a linear time-invariant system. Let $y = H(e^{st})$ be the response of H to the exponential signal e^{st}. Consider the response of H to the time-shifted signal $e^{s(t+\tau)}$ for $\tau \in \mathbb{R}$. By time-invariance

$$H(e^{s(t+\tau)}, t) = H(e^{st}, t+\tau) = y(t+\tau) \qquad \text{for all } t, \tau \in \mathbb{R},$$

and by linearity

$$H(e^{s(t+\tau)}, t) = e^{s\tau} H(e^{st}, t) = e^{s\tau} y(t) \qquad \text{for all } t, \tau \in \mathbb{R}.$$

Combining these equations we obtain

$$y(t+\tau) = e^{s\tau} y(t) \qquad \text{for all } t, \tau \in \mathbb{R}.$$

This equation is satisfied by signals of the form $y(t) = \lambda e^{st}$ where λ is a complex number. That is, the response of H to an exponential signal e^{st} is the same signal e^{st} multiplied by some constant complex number λ. Due to this property exponential signals are called **eigenfunctions** of linear time-invariant systems. The constant λ does not depend on t, but it does usually depend on the complex number s and the system H. To highlight this we will write $\lambda(H, s)$ which, considered as a function of s, is called the **transfer function** of H. Thus, the transfer function satisfies

$$H(e^{st}) = \lambda(H, s) e^{s\tau}. \tag{3.5}$$

We can use these eigenfunctions to better understand the properties of systems modelled by differential equations, such as those in Section 2. As an example, consider the active electrical circuit from Figure 18. In the case that the resistors $R_1 = R_2$, and the capacitor $C_1 = 0$ (an open circuit) the differential equation relating the input voltage x and output voltage y is

$$x = -y - R_1 C_2 D(y).$$

We called this the **active RC** circuit. To simplify notation put $R = R_1$ and $C = C_2$ so that $x = -y - RCD(y)$. In Section 2.2 we were able to solve for the input signal x, given the output signal y. We will now show how to solve for y given x in the special case that x is of the form

$$x = \sum_{\ell=1}^{m} c_\ell e^{s_\ell t}, \tag{3.6}$$

where $c_1, \dots, c_m \in \mathbb{C}$. That is, in the case that x is a linear combination of complex exponential signals.

First observe what occurs when $y = ce^{st}$ is a complex exponential signal with $c \in \mathbb{C}$. We have

$$x = -ce^{st} - cRCse^{st} = -(1 + RCs)ce^{st} = -(1 + RCs)y,$$

and so, x is also a complex exponential signal. We immediately obtain the relationship

$$y = -\frac{1}{1 + RCs} x,$$

that holds whenever y (or equivalently x) is of the form ce^{st} with $c \in \mathbb{C}$. Let H be a system mapping x to y, i.e., such that $y = H(x)$. For now we work under the assumption that H is linear. That this is indeed the case will become clear later. Putting $x = e^{st}$ in the equation above, we find that

$$y = H(x) = H(e^{st}) = -\frac{1}{1 + RCs} e^{st},$$

and so, the transfer function of H is

$$\lambda(H, s) = -\frac{1}{1 + RCs}. \tag{3.7}$$

Now consider when x is a linear combination of complex exponential signals as in (3.6). The output voltage y is

$$\begin{aligned} y = H(x) &= H\left(\sum_{\ell=1}^{m} c_\ell e^{s_\ell t}\right) \\ &= \sum_{\ell=1}^{m} c_\ell H\big(e^{s_\ell t}\big) \qquad \text{(linearity of } H\text{)} \\ &= \sum_{\ell=1}^{m} c_\ell \lambda(H, s_\ell) e^{s_\ell t} \\ &= -\sum_{\ell=1}^{m} \frac{c_\ell e^{s_\ell t}}{1 + RCs_\ell}. \end{aligned} \tag{3.8}$$

Thus, the output signal is also a linear combination of complex exponential signals. The weights in the linear combination are determined by the transfer function.

In Test 3 we used a computer soundcard to pass an approximation of the voltage signal

$$x(t) = \tfrac{1}{3}\sin(2\pi f_1 t) + \tfrac{1}{3}\sin(2\pi f_2 t), \qquad f_1 = 500, f_2 = 1333 \tag{3.9}$$

through the active RC circuit. We are now in a position to derive the output signal y corresponding with this particular input signal. First construct the complex valued signal

$$\begin{aligned} x_a(t) &= \frac{1}{3}\big(j\sin(2\pi f_1 t) + \cos(2\pi f_1 t)\big) + \frac{1}{3}\big(j\sin(2\pi f_2 t) + \cos(2\pi f_2 t)\big) \\ &= \tfrac{1}{3}e^{2\pi f_1 t j} + \tfrac{1}{3}e^{2\pi f_2 t j}, \end{aligned}$$

and observe that the input signal x from (3.9) is the imaginary part of x_a, that is, $x(t) = \mathrm{Im}(x_a(t))$. Suppose x_a is input to the circuit[2]. According to (3.8) the output signal y_a satisfies

$$\begin{aligned} y_a(t) &= H(x_a) \\ &= H\left(\tfrac{1}{3}e^{2\pi f_1 tj} + \tfrac{1}{3}e^{2\pi f_2 tj}\right) \\ &= -\frac{e^{2\pi f_1 tj}}{3 + 6\pi RC f_1 j} - \frac{e^{2\pi f_2 tj}}{3 + 6\pi RC f_2 j}. \end{aligned} \tag{3.10}$$

To extract the desired solution from y_a observe that

$$\begin{aligned} y_a = \mathrm{Re}(y_a) + j\,\mathrm{Im}(y_a) &= H(x_a) \\ &= H\big(\mathrm{Re}(x_a) + j\,\mathrm{Im}(x_a)\big) \\ &= H\big(\mathrm{Re}(x_a)\big) + jH\big(\mathrm{Im}(x_a)\big) \\ &= H\big(\mathrm{Re}(x_a)\big) + jH(x), \end{aligned}$$

and so,

$$y = \mathrm{Im}(y_a) = H(x).$$

That is, the output voltage signal y is the imaginary part of y_a (see Exercise 3.4 for an explicit solution).

3.5 The spectrum

It is often of interest to focus on the transfer function when s is purely imaginary, that is, when $s = j\omega$. In this case the complex exponential signal takes the form

$$e^{j\omega t} = \cos(\omega t) + j\sin(\omega t).$$

This signal is oscillatory when $\omega \neq 0$ and does not decay or explode as $|t| \to \infty$. The function

$$\Lambda(H, f) = \lambda\big(H, j2\pi f\big)$$

is called the **spectrum** of the system H. It follows from (3.5) that the response of the system to the complex exponential signal $e^{j2\pi ft}$ satisfies

$$H(e^{j2\pi ft}) = \lambda(H, j2\pi f)e^{j2\pi ft} = \Lambda(H, f)e^{j2\pi ft}, \qquad f \in \mathbb{R}.$$

It is of interest to consider the **magnitude spectrum** $|\Lambda(H, f)|$ and the **phase spectrum** $\angle\Lambda(H, f)$ separately. The notation $\angle$ denotes the **argument** (or **phase**) of a complex number. We have,

$$\Lambda(H, f) = |\Lambda(H, f)|\, e^{j\angle\Lambda(H,f)},$$

and so,

$$H(e^{j2\pi ft}) = |\Lambda(H, f)|\, e^{j(2\pi ft + \angle\Lambda(H,f))}.$$

[2]In practice, we cannot input a complex signal to any electrical circuit. However, it is instructive to temporarily pretend that we can, and to carry the equations through.

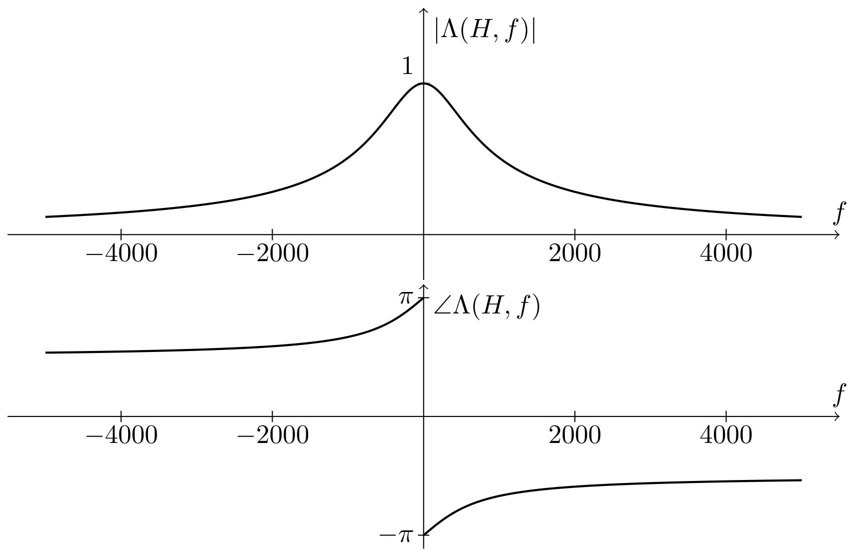

Figure 28: Magnitude spectrum (top) and phase spectrum (bottom) of the active RC circuit with $R = 27 \times 10^3$ and $C = 10 \times 10^{-9}$.

By taking real and imaginary parts we obtain the pair of real valued solutions

$$\begin{aligned} H\big(\cos(2\pi f t)\big) &= |\Lambda(H,f)| \cos\big(2\pi f t + \angle\Lambda(H,f)\big), \\ H\big(\sin(2\pi f t)\big) &= |\Lambda(H,f)| \sin\big(2\pi f t + \angle\Lambda(H,f)\big). \end{aligned} \tag{3.11}$$

Consider again the active RC circuit with H the system mapping the input voltage x to the output voltage y. According to (3.7) the spectrum of H is

$$\Lambda(H,f) = -\frac{1}{1 + 2\pi RCfj}.$$

The magnitude and phase spectrum is

$$|\Lambda(H,f)| = \big(1 + 4\pi^2 R^2 C^2 f^2\big)^{-\frac{1}{2}}, \qquad \angle\Lambda(H,f) = \operatorname{atan}\big(2\pi RCf\big) + \pi.$$

The magnitude and phase spectrum are plotted in Figure 28. Observe from the plot of the magnitude spectrum that a low frequency sinusoidal signal, say 100Hz or less, input to the RC circuit results in a sinusoidal output signal with the same frequency and approximately the same amplitude. However, a high frequency sinusoidal signal, say greater than 1000Hz, input to the RC circuit results in a sinusoidal output signal with the same frequency, but small amplitude. For this reason RC circuits are called **low pass filters**.

Test 4 (**Spectrum of the active RC circuit**) We test the hypothesis that the active RC circuit satisfies (3.11). To do this sinusoidal signals at varying frequencies of the form

$$x_k(t) = \sin(2\pi f_k t), \qquad f_k = 110 \times 2^{k/2}, \quad k = 0, 1, \dots, 12$$

are input to the active RC circuit constructed as in Test 3 with $R = R_1 = 27\text{k}\Omega$ and $C = C_2 = 10\text{nF}$. In view of (3.11) the expected output signals are of the form

$$y_k(t) = |\Lambda(H, f_k)| \sin\big(2\pi f_k t + \angle\Lambda(H, f_k)\big), \qquad k = 0, 1, \dots, 21.$$

For any positive integer M the energy of the periodic transmitted signal x_k over any interval of length $T = M/f_k$ (an interval containing M periods) is

$$\text{energy}(x_k) = \int_0^T \sin^2(2\pi f_k t) dt = \frac{1}{2}\int_0^T 1 - \cos(4\pi f_k t) dt = \frac{T}{2} = \frac{M}{2f_k}.$$

The energy of the output signal y_k over the same interval is

$$\text{energy}(y_k) = |\Lambda(H, f_k)|^2 \,\text{energy}(x_k) = \frac{\text{energy}(x_k)}{1 + 4\pi^2 R^2 C^2 f_k^2}. \tag{3.12}$$

We see that the square of the magnitude spectrum relates the energy of the input and output signals. We test this relationship.

Using the soundcard the signals x_k for each $k = 0, \dots, 21$ are input to the circuit. Reconstructions of the input signal $\tilde{x}_k$ and the output signal $\tilde{y}_k$ are constructed from samples $x_{k,1}, \dots, x_{k,L}$ and $y_{k,1}, \dots, y_{k,L}$ in a similar manner to (1.8) and (1.6) where L is the number of samples obtained by the soundcard. The energy of the reconstructed input signal $\tilde{x}_k$ is

$$\begin{aligned}
\|\tilde{x}_k\|_2 &= \int_{-\infty}^{\infty} \left| \sum_{\ell=1}^{L} x_{k,\ell}\,\text{sinc}(F_s t - \ell) \right|^2 dt \\
&= \int_{-\infty}^{\infty} \sum_{\ell=1}^{L}\sum_{m=1}^{L} x_{k,\ell} x_{k,m}\,\text{sinc}(F_s t - \ell)\,\text{sinc}(F_s t - m) dt \\
&= \sum_{\ell=1}^{L}\sum_{m=1}^{L} x_{k,\ell} x_{k,m} \int_{-\infty}^{\infty} \text{sinc}(F_s t - \ell)\,\text{sinc}(F_s t - m) dt \\
&= \frac{1}{F_s} \sum_{\ell=1}^{L} x_{k,\ell}^2
\end{aligned}$$

where, on the last line we use the fact that sinc and its time shifts by a nonzero integer $T_m(\text{sinc})$ are **orthogonal** (Exercise 5.1). That is,

$$\int_{-\infty}^{\infty} \text{sinc}(t)\,\text{sinc}(t - m) dt = \begin{cases} 1 & m = 0 \\ 0 & m \neq 0. \end{cases} \tag{3.13}$$

Similarly, the energy of the reconstructed output signal $\tilde{y}_k$ is

$$\|\tilde{y}_k\|_2 = \frac{1}{F_s}\sum_{\ell=1}^{L} y_{k,\ell}^2.$$

So, to compute the energy of the reconstructed signals it suffices to sum the squares of the samples and divide by the sample rate F_s. In view of 3.12, we expect the approximate relationship

$$\frac{\|\tilde{y}_k\|_2}{\|\tilde{x}_k\|_2} \approx |\Lambda(H, f_k)|^2 = \frac{1}{1 + 4\pi^2 R^2 C^2 f_k^2}. \tag{3.14}$$

Each signal x_k is played for a period of approximately 1 second and approximately $L \approx F_s = 44100$ samples are obtained. On the soundcard hardware used for this test samples near the beginning and end of playback are distorted. This appears to be an unavoidable feature of the soundcard. To alleviate this we discard the first $A - 1 = 9999$ samples and use only the $B = 8820$ samples that follow (corresponding to 200ms of signal). In view of (3.14), we expect the relationship

$$\sqrt{\frac{\sum_{\ell=A}^{A+B} y_{k,\ell}^2}{\sum_{\ell=A}^{A+B} x_{k,\ell}^2}} \approx |\Lambda(H, f)| = \sqrt{\frac{1}{1 + 4\pi^2 R^2 C^2 f_k^2}}.$$

Figure 29 displays a plot of the hypothesised spectrum $|\Lambda(H, f)|$ (solid line) and also the spectrum measured using the left hand side of the approximate equation above (dots). The measurements are close to the hypothesised spectrum, but are consistently a small amount larger. This is similar to the effect observed in Test 2. The amplifier appears to produce a slightly larger output voltage than expected. This could be due to inaccuracies in the components used, and also due to our assumption of an ideal operational amplifier.

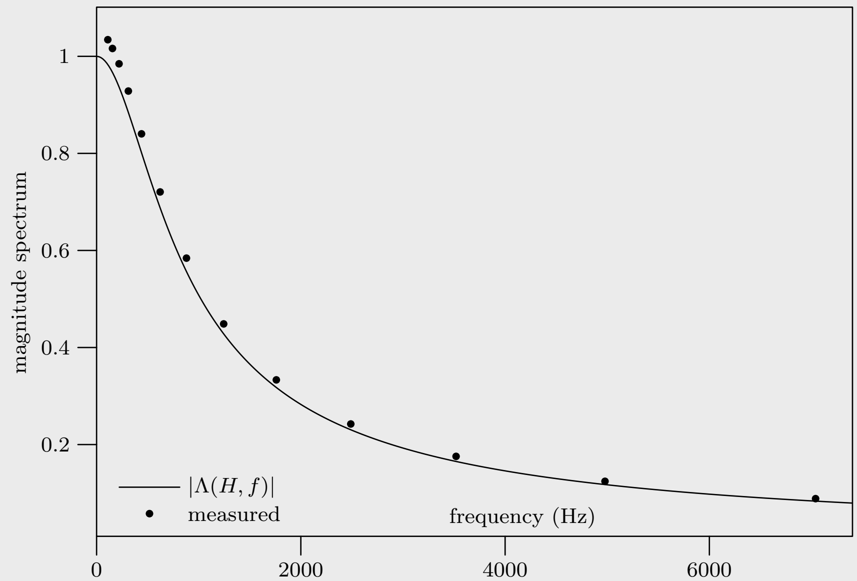

Figure 29: Plot of the hypothesised spectrum $|\Lambda(H, f)|$ (solid line) and the measured spectrum (dots).

3.6 Exercises

3.1. Show that convolution distributes with addition and commutes with scalar multiplication, that is, show that $a(x * w) + b(y * w) = (ax + by) * w$.

3.2. Show that convolution is associative. That is, if x, y, z are signals then $x * (y * z) = (x * y) * z$.

3.3. Show that a regular system is stable if and only if its impulse response is absolutely integrable.

3.4. Find an explicit formula for the imaginary part of the signal y_a from (3.10).

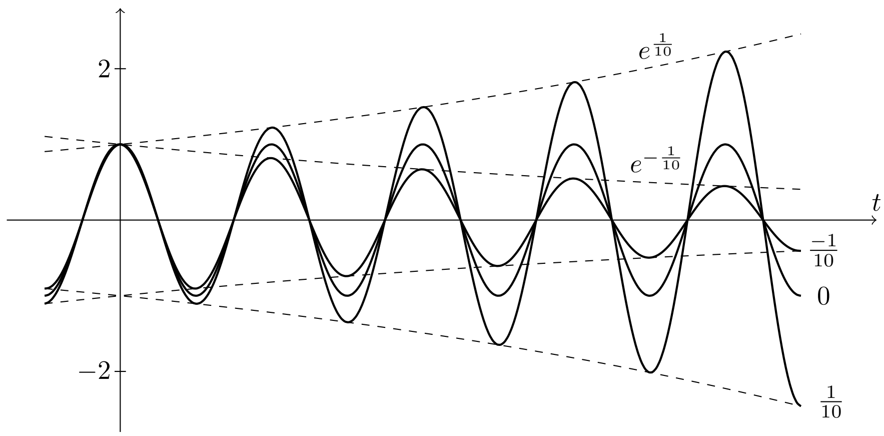

Figure 30: The function $\cos(\pi t)e^{\sigma t}$ for $\sigma = -\frac{1}{10}, 0, \frac{1}{10}$.

4 The Laplace transform

Let $x\colon \mathbb{R} \to \mathbb{C}$ be a complex valued function of the real line (a signal). The integral

$$\mathcal{L}(x) = \int_{-\infty}^{\infty} x(t)e^{-st}dt, \tag{4.1}$$

when it exists, is called the **Laplace transform** of x. The Laplace transform is a function of the complex parameter s, and if we need to indicate this we write $\mathcal{L}(x)(s)$ or $\mathcal{L}(x,s)$. The Laplace transform does not necessarily exist for all values of $s \in \mathbb{C}$. Let R be the set of real numbers such that $x(t)e^{-\sigma t}$ is absolutely integrable if and only if $\sigma \in R$, that is

$$\int_{-\infty}^{\infty} |x(t)|\, e^{-\sigma t} dt \qquad \text{exists if and only if } \sigma \in R.$$

In this case, the Laplace transform $\mathcal{L}(x,s)$ exists for all s with real part satisfying $\mathrm{Re}(s) \in R$ because

$$\mathcal{L}(x) = \int_{-\infty}^{\infty} x(t)e^{-st}dt \leq \int_{-\infty}^{\infty} |x(t)|\, e^{-\mathrm{Re}(s)t}dt < \infty.$$

The subset of the complex plane with real part from R is called the **region of convergence** (ROC) of the signal x.

For example, the Laplace transform of the right sided signal $e^{\alpha t}u(t)$ is

$$\begin{aligned}
\mathcal{L}(e^{\alpha t}u(t)) &= \int_{-\infty}^{\infty} e^{\alpha t}e^{-st}u(t)dt \\
&= \int_{0}^{\infty} e^{(\alpha - s)t}dt \\
&= \lim_{t\to\infty} \frac{e^{(\alpha-s)t}}{\alpha - s} - \frac{1}{\alpha}.
\end{aligned}$$

This transform exists for all s with $\text{Re}(s) > \text{Re}(\alpha)$. The region of convergence of the function $e^{\alpha t}u(t)$ is the subset of the complex plane with real part greater than $\text{Re}(\alpha)$. Figure 31 shows the region of convergence when $\text{Re}(\alpha) = -2$. Now consider the left sided signal $e^{\beta t}u(-t)$ with Laplace transform

$$\mathcal{L}(e^{\beta t}u(-t)) = \lim_{t\to-\infty} \frac{e^{(\beta-s)t}}{\alpha - s} + \frac{1}{\alpha}$$

that exists only when $\text{Re}(s) < \text{Re}(\beta)$.

The signal $ae^{\alpha t}u(t) + be^{\beta t}u(-t)$ has Laplace transform

$$\begin{aligned}
\mathcal{L}\big(ae^{\alpha t}u(t) + be^{\beta t}u(-t)\big) &= \int_{-\infty}^{\infty} \big(ae^{\alpha t}u(t) + be^{\beta t}u(-t)\big)e^{-st}dt \\
&= a\int_{-\infty}^{\infty} e^{\alpha t}u(t)e^{-st}dt + b\int_{-\infty}^{\infty} e^{\beta t}u(-t)\big)e^{-st}dt \\
&= a\mathcal{L}(e^{\alpha t}u(t)) + b\mathcal{L}(e^{\beta t}u(-t))
\end{aligned}$$

that exists only when $\text{Re}(\alpha) < \text{Re}(s) < \text{Re}(\beta)$. The corresponding ROC is shown in Figure 31 when $\text{Re}(\alpha) = -2$ and $\text{Re}(\beta) = 3$. In the previous equation we have discovered that the Laplace transform is **linear**, that is, for signals x and y and constants a and b,

$$\mathcal{L}(ax + by) = a\mathcal{L}(x) + b\mathcal{L}(y). \tag{4.2}$$

In words: the Laplace transform of a linear combination of signals is the same linear combination of the Laplace transforms of those signals.

In the previous example the Laplace transform does not exist for any s if $\text{Re}(\alpha) \geq \text{Re}(\beta)$, and the region of convergence is correspondingly the empty set. Other signals also have this property. For example $x(t) = 1$ does not have a Laplace transform because

$$\mathcal{L}(1) = \int_{\infty}^{\infty} e^{-st}dt = \frac{1}{s}\lim_{t\to-\infty} e^{-st} - \frac{1}{s}\lim_{t\to\infty} e^{-st},$$

and the left hand limit exists only when $\text{Re}(s) < 0$ while the right hand limit exists only when $\text{Re}(s) > 0$. We can similarly show that any periodic signal does not have a Laplace transform (Exercise 4.5).

As a final example, consider the rectangular pulse

$$\Pi(t) = \begin{cases} 1 & -\frac{1}{2} < t \leq \frac{1}{2} \\ 0 & \text{otherwise.} \end{cases}$$

Its Laplace transform is

$$\mathcal{L}(\Pi) = \int_{-\infty}^{\infty} \Pi(t)e^{-st}dt = \int_{-1/2}^{1/2} e^{-st}dt = \frac{e^{s/2} - e^{-s/2}}{s}, \tag{4.3}$$

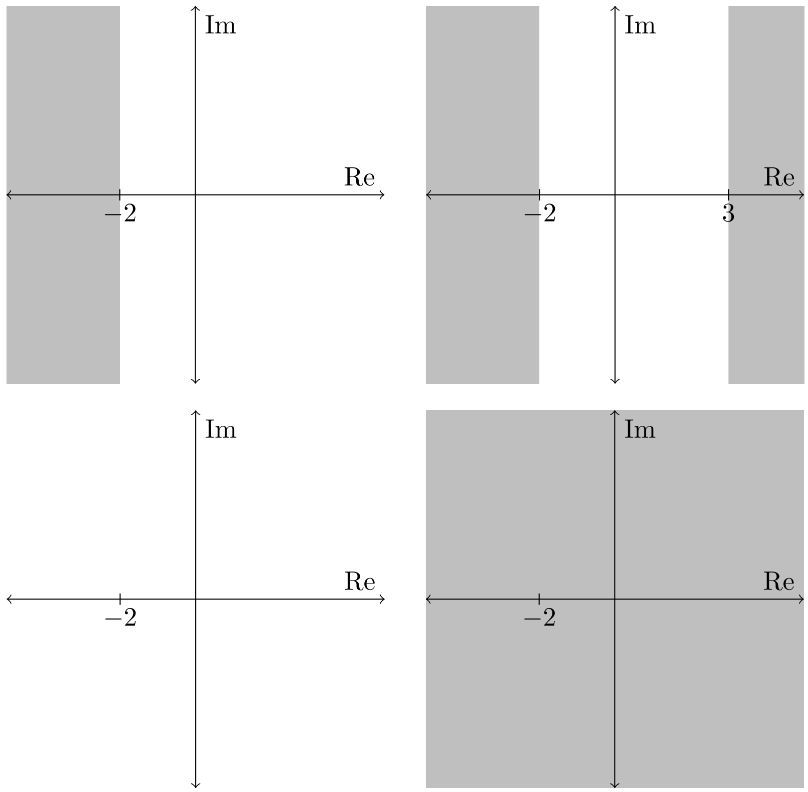

Figure 31: Regions of convergence (unshaded) for the signal $e^{-2t}u(t)$ (top left), the signal $e^{-2t}u(t) + e^{3t}u(-t)$ (top right), the rectangular pulse Π (bottom left), and the constant signal $x(t) = 1$ (bottom right).

and this transform exists for all $s \in \mathbb{C}$. The region of convergence of the rectangular pulse Π is the entire complex plane. The examples just given exhibit all the possible types of regions of convergence. The region of convergence is either the entire complex plane, a left or right half plane, a vertical strip, or the empty set.

Given the Laplace transform $\mathcal{L}(x)$ the signal x can be recovered by the **inverse Laplace transform**

$$x(t) = \mathcal{L}^{-1}(x) = \frac{1}{2\pi j} \lim_{\omega \to \infty} \int_{\sigma - j\omega}^{\sigma - j\omega} \mathcal{L}(x, s) e^{st} ds,$$

where σ is a real number that is inside the region of convergence of x. Solving the integral above typically requires a special type of integration called **contour integration** that we will not consider here [Stewart and Tall, 2004]. For our purposes, and for many engineering purposes, it suffices to remember only the following Laplace transform pair

$$\mathcal{L}\big(t^n u(t)\big) = \frac{n!}{s^{n+1}} \qquad \text{Re}(s) > 0,$$

where n is an integer greater than zero (Exercise 4.6). Combining this with the **frequency shift rule**,

$$\begin{aligned}\mathcal{L}\big(e^{\alpha t}x(t), s\big) &= \int_{-\infty}^{\infty} e^{\alpha t}x(t)e^{-st}dt \\ &= \int_{-\infty}^{\infty} x(t)e^{-(s-\alpha)t}dt \\ &= \mathcal{L}(x, s-\alpha) \qquad \text{Re}(s-\alpha) \in R,\end{aligned}$$

where R is the region of convergence of x, we obtain the transform pair

$$\mathcal{L}\big(t^n e^{\alpha t} u(t)\big) = \mathcal{L}\big(t^n u(t), s-\alpha\big) = \frac{n!}{(s-\alpha)^{n+1}} \qquad \text{Re}(s) > \text{Re}(\alpha), \tag{4.4}$$

where n is an integer greater than zero. This is the only Laplace transform pair we require here.

4.1 The transfer function and the Laplace transform

Our purpose for introducing the Laplace transform is to study the response of a linear time-invariant system H to exponential signals of the form e^{st}. Recall from Section 3.4 that exponential signals are **eigenfunctions** of linear time-invariant systems. That is, for each $s \in \mathbb{C}$, the response of H to e^{st} is λe^{st} where $\lambda \in \mathbb{C}$ is a constant that does not depend on t, but may depend on s and the system H. To highlight this dependence on H and s we write $\lambda(H, s)$ which, considered as a function of s, is called the **transfer function** of H. For a given system H, we would like to understand how $\lambda(H, s)$ behaves as s changes. Let us first assume that H is a regular system with impulse response h. In this case,

$$\begin{aligned}H(e^{st}, t) = e^{st}\lambda(H, s) &= h * e^{-st} \\ &= \int_{-\infty}^{\infty} h(\tau)e^{-s(t-\tau)}d\tau \\ &= e^{st}\int_{-\infty}^{\infty} h(\tau)e^{-s\tau}d\tau \\ &= e^{st}\mathcal{L}(h, s),\end{aligned}$$

and so, $\lambda(H) = \mathcal{L}(h)$. That is, the transfer function of a regular system is precisely the Laplace transform of the impulse response. The region of convergence of the impulse response describes the set of complex exponential signals e^{st} that can be input to the system and we refer to this as the region of convergence of the *system*. In this way, both signals and systems have regions of convergence.

The transfer functions of the time-shifter and differentiator can be obtained by inspection. For the time-shifter

$$T_\tau(e^{st}) = e^{s(t-\tau)} = e^{-s\tau}e^{st} \qquad \text{and so} \qquad \lambda(T_\tau, s) = e^{-s\tau}. \tag{4.5}$$

The region of convergence is the whole complex plane $s \in \mathbb{C}$. For the special case of the identity system T_0 we obtain $\lambda(T_0, s) = 1$. For the differentiator

$$D(e^{st}) = \frac{d}{dt}e^{st} = se^{st} \qquad \text{and so} \qquad \lambda(D, s) = s.$$

The region of convergence is the whole complex plane $s \in \mathbb{C}$. More generally, for the kth differentiator

$$D^k(e^{st}) = \frac{d^k}{dt^k}e^{st} = s^k e^{st} \qquad \text{and so} \qquad \lambda(D^k, s) = s^k. \tag{4.6}$$

The region of convergence is again the whole complex plane. These results motivate assigning the following Laplace transforms to the delta "function" and its derivatives

$$\mathcal{L}(\delta, s) = 1, \qquad \mathcal{L}(\delta^k, s) = s^k.$$

These conventions are common in the engineering literature Oppenheim et al. [1996].

4.1.1 The transfer function of a composition of systems

Let H be the system constructed by composing systems H_1 and H_2, that is, $H(x) = H_1\big(H_2(x)\big)$. Let $R_1 \subseteq \mathbb{C}$ and $R_2 \subseteq \mathbb{C}$ be the regions on convergence of H_1 and H_2. The response of H to the signal e^{st} is

$$H(e^{st}) = H_1\big(H_2(e^{st})\big) = H_1\big(\lambda(H_2)e^{st}\big) = \lambda(H_1)\lambda(H_2)e^{st},$$

and so, the transfer function of H is given by the multiplication of the transfer functions of H_1 and H_2. The region of convergence of H is the intersection of R_1 and R_2. That is,

$$\lambda(H) = \lambda(H_1)\lambda(H_2), \qquad s \in R_1 \cap R_2. \tag{4.7}$$

4.1.2 The convolution theorem

We showed in Section 3.3 that if H_1 and H_2 are regular systems with impulse responses h_1 and h_2, then the impulse of the system $H(x) = H_1\big(H_2(x)\big)$ is given by the convolution $h = h_1 * h_2$. Because,

$$\lambda(H) = \mathcal{L}(h) \qquad \lambda(H_1) = \mathcal{L}(h_1) \qquad \lambda(H_1) = \mathcal{L}(h_1),$$

and using (4.7), we obtain,

$$\mathcal{L}(h_1 * h_2) = \mathcal{L}(h) = \lambda(H) = \lambda(H_1)\lambda(H_2) = \mathcal{L}(h_1)\mathcal{L}(h_2), \qquad s \in R_1 \cap R_2. \tag{4.8}$$

This is called the **convolution theorem** and goes by the phrase: "Convolution in the time domain is multiplication in the Laplace/Fourier/Frequency domain".

4.1.3 The Laplace transform of an output signal

Let H be a regular system with impulse response h and let $y = H(x) = h * x$ be the response of H to input signal x. Using the convolution theorem, the Laplace transform of the output signal y is

$$\mathcal{L}(y) = \mathcal{L}(h)\mathcal{L}(x) = \lambda(H)\mathcal{L}(x), \qquad s \in R \cap R_x, \tag{4.9}$$

where R is the region of convergence of the system H and R_x is the region of convergence of the input signal x. Thus, the Laplace transform of the output signal $y = H(x)$ is the transfer function of the system H multiplied by the Laplace transform of the input signal x. This result holds even when H is not regular, for example when H is a time-shifter or a differentiator (Exercise 4.9).

4.2 Solving differential equations

Assume we have a system modelled by a differential equation of the form

$$\sum_{\ell=0}^{m} a_\ell D^\ell(x) = \sum_{\ell=0}^{k} b_\ell D^\ell(y), \tag{4.10}$$

where x and y are signals. Taking Laplace transforms of both sides of this equation,

$$\begin{aligned}
\mathcal{L}\left(\sum_{\ell=0}^{m} a_\ell D^\ell(x)\right) &= \mathcal{L}\left(\sum_{\ell=0}^{k} b_\ell D^\ell(y)\right) \\
\sum_{\ell=0}^{m} a_\ell \mathcal{L}\big(D^\ell(x)\big) &= \sum_{\ell=0}^{k} b_\ell \mathcal{L}\big(D^\ell(x)\big) && \text{(linearity (4.2))} \\
\sum_{\ell=0}^{m} a_\ell \lambda(D^\ell)\mathcal{L}(x) &= \sum_{\ell=0}^{k} b_\ell \lambda(D^\ell)\mathcal{L}(y) && \text{(using (4.9))} \\
\sum_{\ell=0}^{m} a_\ell s^\ell \mathcal{L}(x) &= \sum_{\ell=0}^{k} b_\ell s^\ell \mathcal{L}(y). && \text{(since } \lambda(D^\ell) = s^\ell \text{ by (4.6))}
\end{aligned}$$

We have obtained an equation relating the Laplace transforms of x and y,

$$\mathcal{L}(x)(a_0 + a_1 s + \dots a_m s^m) = \mathcal{L}(y)(b_0 + b_1 s + \dots b_k s^k).$$

Rearranging this equation we obtain,

$$\mathcal{L}(y) = \frac{a_0 + a_1 s + \dots a_m s^m}{b_0 + b_1 s + \dots b_k s^k} \mathcal{L}(x).$$

Given the input signal x and its Laplace transform $\mathcal{L}(x)$ we can find the output signal y (provided it exists) by applying the inverse Laplace transform to the right hand side of the equation above.

Let H be a system such that $y = H(x)$ whenever x and y satisfy the differential equation (4.10). According to (4.9) the transfer function of H is

$$\lambda(H) = \frac{\mathcal{L}(y)}{\mathcal{L}(x)} = \frac{a_0 + a_1 s + \dots a_m s^m}{b_0 + b_1 s + \dots b_k s^k}.$$

Properties of H can be obtained by inspecting this transfer function. For example, the impulse response of H (if is exists) can be obtained by applying the inverse Laplace transform.

4.3 First order systems

Recall the passive electric RC circuit from Figure 11. The differential equation modelling this circuit is (2.1)

$$x = y + RCD(y),$$

where x is the input voltage signal, y is the voltage over the capacitor and R and C are the resistance and capacitance. The RC circuit is an example of a **first order system**. Let H be a system mapping the input voltage x to output voltage y. We will discover the impulse response of H. Taking the Laplace transform on both sides of the differential equation gives

$$\mathcal{L}(x) = (1 + RCs)\mathcal{L}(y),$$

and it follows that the transfer function of H is

$$\lambda(H) = \frac{\mathcal{L}(y)}{\mathcal{L}(x)} = \frac{1}{1 + RCs} = \frac{r}{r + s},$$

where $r = \frac{1}{RC}$. The value $\frac{1}{r} = RC$ is called the **time constant**. The impulse response of H is given by the inverse of this Laplace transform. There are two signals with Laplace transform $\frac{r}{r+s}$: the right sided signal $re^{-rt}u(t)$ with region of convergence $\mathrm{Re}(s) > -r$, and the left sided signal $-re^{-rt}u(-t)$ with region of convergence $\mathrm{Re}(s) < -r$. The RC circuit (and infact all phyically realisable systems) are expected to be causal. For this reaon, the left sided signal $-re^{-rt}u(-t)$ cannot be the impulse response of H. The impulse response is the right sided signal

$$h(t) = re^{-rt}u(t).$$

Given an input voltage signal x we can now find the corresponding output signal $y = H(x)$ by convolving x with the impulse response h. That is,

$$y = H(x) = h * x = \int_{-\infty}^{\infty} re^{-r\tau}u(\tau)x(t-\tau)d\tau = r\int_{0}^{\infty} e^{-r\tau}x(t-\tau)d\tau.$$

The output signal y exists whenever this convolution exists.

If $r \geq 0$ the impulse response is absolutely summable, that is,

$$\begin{aligned}\|h\|_1 &= \int_{-\infty}^{\infty} \left| re^{-rt} u(t) \right| dt \\ &= r \int_{0}^{\infty} e^{-rt} dt \\ &= r - r \lim_{t \to \infty} e^{-rt} = r,\end{aligned}$$

and the system is stable (Exercise 3.3). However, if $r < 0$ the impulse response is not absolutely summable, and the system is not stable. Figure 33 shows the impulse response when $r = -\frac{1}{5}, -\frac{1}{3}, -\frac{1}{2}, -\frac{1}{2}, 1, 2$. In a electrical RC circuit the resistance R and capacitance C are always positive and $r = \frac{1}{RC}$ is positive. For this reason, passive RC circuits are always stable.

From (3.4), the step response $H(u)$ is given by applying the integrator I_∞ to the impulse response,

$$H(u) = I_\infty(h) = \int_{-\infty}^{t} \tau e^{-r\tau} u(\tau) d\tau = \tau \int_{0}^{t} e^{-r\tau} d\tau = 1 - e^{-rt}. \tag{4.11}$$

This step response in plotted in Figure 33.

Test 5 (**The active RC circuit again**) In this test we repeat the experiment with the active RC circuit from Test 3 with resistors $R = R_1 = R_2 = 27\text{k}\Omega$ and capacitors $C = C_2 = 10\text{nF}$. In Test 3 we applied the differential equation (2.8) to the reconstructed output signal $\tilde{y}$ and asserted that the resulting signal was close to the reconstructed input signal $\tilde{x}$. In this test we instead convolve the input signal $\tilde{x}$ with the impulse response

$$h = -\tfrac{1}{RC} e^{-t/RC} = -re^{-rt}, \qquad r = \tfrac{1}{RC} = \frac{100000}{27},$$

and assert that the resulting signal is close to the output signal $\tilde{y}$. That is, we test the expected relationship

$$\tilde{y} \approx h * \tilde{x} = -\int_{-\infty}^{\infty} re^{-r\tau} u(\tau) \tilde{x}(t - \tau) d\tau = -r \int_{0}^{\infty} e^{-r\tau} \tilde{x}(t - \tau) d\tau.$$

From (1.8),

$$\begin{aligned}\tilde{y}(t) &\approx -r \int_{0}^{\infty} e^{-r\tau} \sum_{\ell=1}^{L} x_\ell \operatorname{sinc}(F_s t - F_s \tau - \ell) d\tau \\ &= -r \sum_{\ell=1}^{L} x_\ell \int_{0}^{\infty} e^{-r\tau} \operatorname{sinc}(F_s t - F_s \tau - \ell) d\tau \\ &= -r \sum_{\ell=1}^{L} x_\ell f(F_s t - \ell),\end{aligned}$$

where the function

$$f(t) = \int_0^\infty e^{-r\tau} \operatorname{sinc}(t - F_s\tau) d\tau.$$

An approximation of $f(t)$ is made using the trapezoidal sum

$$f(t) \approx \frac{K}{2N} \left(g(0) + g(K) + 2 \sum_{n=1}^{N-1} g(\Delta n) \right),$$

where $g(\tau) = e^{-r\tau} \operatorname{sinc}(t - F_s\tau)$, and

$$K = -RC \log\left(10^{-3}\right), \qquad N = \lceil 10F_sK \rceil, \qquad \Delta = K/N.$$

Figure 32 plots the input signal $\tilde{x}$, output signal $\tilde{y}$ and hypothesised output signal $h * \tilde{x}$ over a 4ms window.

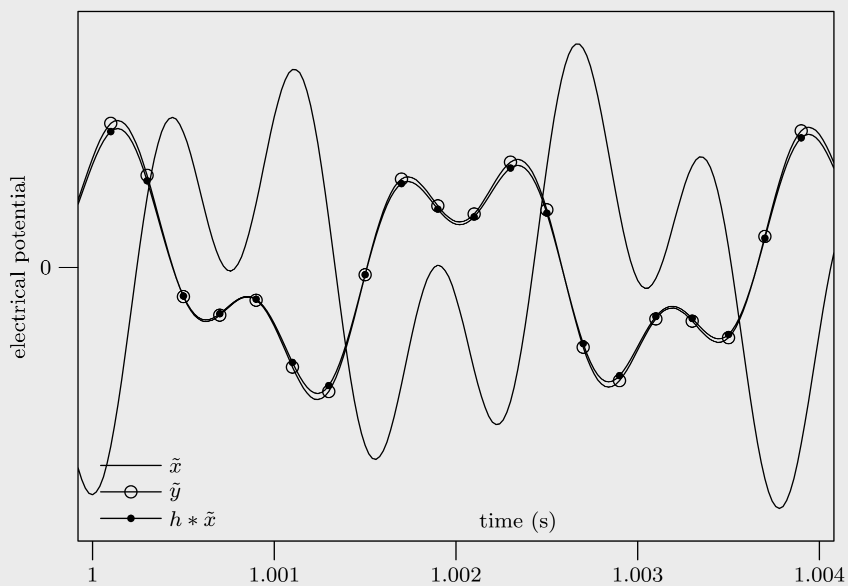

Figure 32: Plot of reconstructed input signal $\tilde{x}$ (solid line), output signal $\tilde{y}$ (solid line with circle), and hypothesised output signal $h * \tilde{x}$ (solid line with dot).

4.4 Second order systems

Consider the mass, spring, damper system from Figure 12 and described by the equation

$$f = Kp + BD(p) + MD^2(p), \tag{4.12}$$

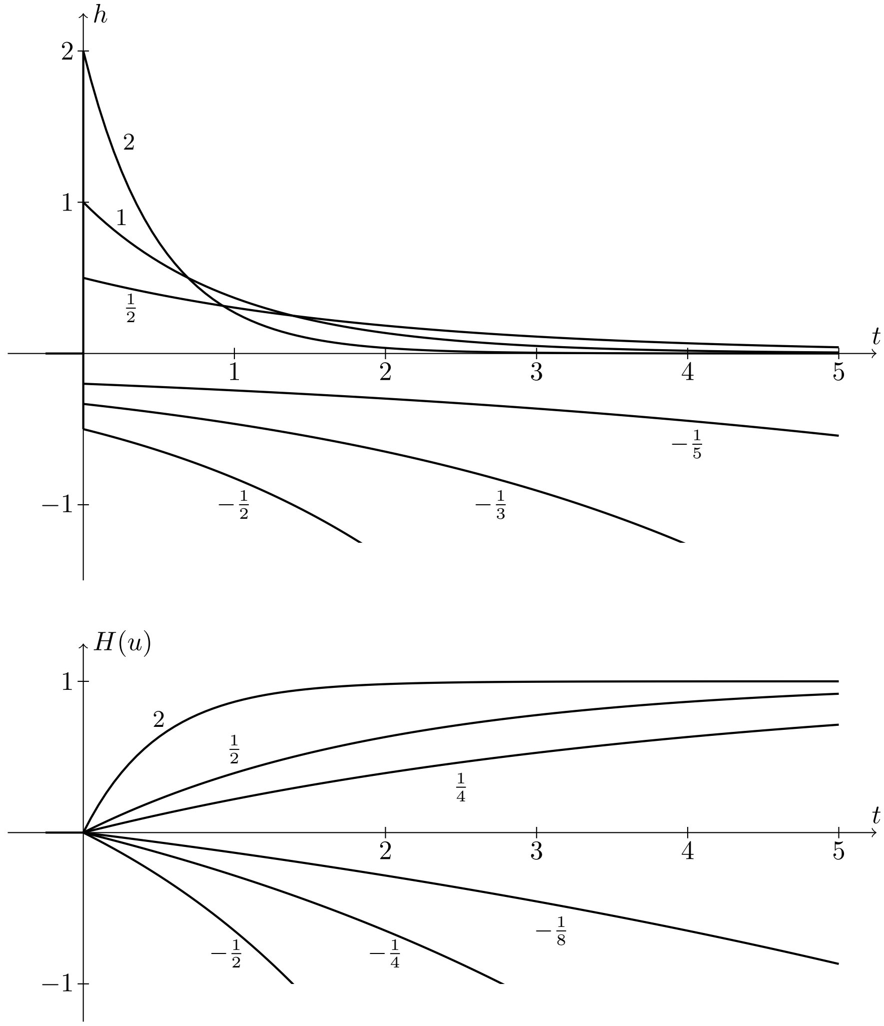

Figure 33: Top: impulse response of a first order system with $r = -\frac{1}{2}$, $-\frac{1}{3}$, $-\frac{1}{5}$, $\frac{1}{2}$, 1, 2. Bottom: step response of a first order system with $r = -\frac{1}{2}$, $-\frac{1}{4}$, $-\frac{1}{8}$, $\frac{1}{4}$, $\frac{1}{2}$, 2.

where f is the force applied to the mass M and p is the position of the mass and K and B are the spring and damping coefficients. In Section 2 we were able to find the force f corresponding with a given position signal p. Let H be a system mapping f to p, that is, such that $p = H(f)$. We will find the impulse response of H. Taking Laplace transforms on both sides of the differential equation gives

$$\mathcal{L}(f) = (K + Bs + Ms^2)\mathcal{L}(p).$$

Rearranging gives the transfer function of H,

$$\lambda(H) = \frac{\mathcal{L}(p)}{\mathcal{L}(f)} = \frac{1}{K + Bs + Ms^2}.$$

We can invert this Laplace transform to obtain the impulse response. There are three cases to consider, depending on whether the quadratic $K + Bs + Ms^2$ has two distinct real roots, is irreducible (does not have real roots), or has two identical real roots.

Case 1: (Distinct real roots) In this case, the roots are

$$\beta - \alpha, \qquad -\beta - \alpha,$$

where

$$\alpha = \frac{B}{2M}, \qquad \beta = \frac{\sqrt{B^2 - 4KM}}{2M}$$

and $B^2 - 4KM > 0$. By a partial fraction expansion (Exercise 4.2)

$$\begin{aligned}\lambda(H) &= \frac{1}{M(s-\beta+\alpha)(s+\beta+\alpha)} \\ &= \frac{1}{2\beta M}\left(\frac{1}{s-\beta+\alpha} - \frac{1}{s+\beta+\alpha}\right).\end{aligned}$$

From (4.4), we obtain the transform pairs

$$\mathcal{L}(e^{(\beta-\alpha)t}u(t)) = \frac{1}{s-\beta+\alpha}, \qquad \mathcal{L}(e^{-(\beta+\alpha)t}u(t)) = \frac{1}{s+\beta+\alpha}.$$

As in Section 4.3, other signals with these Laplace transforms are discarded becuase they do not lead to an impulse response that is zero for $t < 0$. That is, they do not lead to a causal system H. The impulse response of H is now

$$h(t) = \frac{1}{2\beta M}u(t)e^{-\alpha t}\big(e^{\beta t} - e^{-\beta t}\big).$$

This is a sum of the impulse response of two first order systems.

Case 2: (Distinct imaginary roots) The solution is as in the previous case, but now $4KM - B^2 > 0$ and β is imaginary. Put $\theta = \beta/j$ so that

$$e^{\beta t} - e^{-\beta t} = e^{j\theta t} - e^{-j\theta t} = 2j\sin(\theta t),$$

and the impulse response of H is

$$h(t) = \frac{1}{\theta M}u(t)e^{-\alpha t}\sin(\theta t).$$

Case 3: (Identical roots) In this case, the two roots are equal to $-\alpha$ and

$$\lambda(H) = \frac{1}{M(s+\alpha)^2}.$$

From (4.4) we obtain the transform pair

$$\mathcal{L}\big(te^{-\alpha t}u(t)\big) = \frac{1}{(s+\alpha)^2},$$

and this is the only signal with this Laplace transform that leads to a causal impulse response. The impulse response of H is thus

$$h(t) = \frac{1}{M}te^{-\alpha t}u(t).$$

A second order system is called **overdamped** when there are two distinct real roots, **underdamped** when their are two distinct imaginary roots, and **critically damped** when the roots are identical. The different types of impulse responses for are plotted in Figure 34.

With no damping (i.e. damping coefficient $B = 0$) the roots are of the form $\pm\beta$ and have no real part. In this case, the impulse response is

$$h(t) = \frac{1}{\theta M}u(t)\sin(\theta t),$$

where $\theta = \beta/j = \sqrt{KM}$ is called the **natural frequency** of the second order system. This impulse reponse ocillates for all $t > 0$ without decay or explosion. Two identical roots occur when the damping coefficient $B = \sqrt{4KM}$, and this is sometimes called the **critical damping coefficient**.

The impulse response of a second order system is asbolutely integrable, and the system is correspondingly stable, when $\alpha = \frac{B}{2M} > 0$, but not when $\alpha \leq 0$. For the mass spring damper system both the mass M and damping coefficients B are positive, and so, mass spring dampers are always stable.

4.5 Poles, zeros, and stability

As discussed in Section 4.2 the transfer function of a system described by a linear differential equation with constant coefficients is of the form

$$\lambda(H) = \frac{\mathcal{L}(y)}{\mathcal{L}(x)} = \frac{a_0 + a_1 s + \dots a_m s^m}{b_0 + b_1 s + \dots b_m s^m}.$$

Factorising the polynomials on the numerator and denominator we obtain

$$\lambda(H) = \frac{(s-\alpha_0)(s-\alpha_1)\cdots(s-\alpha_m)}{(s-\beta_0)(s-\beta_1)\cdots(s-\beta_k)},$$

where $\alpha_0, \dots, \alpha_m$ are the roots of the numerator polynomial $a_0 + a_1 s + \cdots + a_m s^m$, and $\beta_0, \dots, \beta_m$ are the roots of the denominator polynomial $b_0 + b_1 s + \cdots + b_k s^k$.

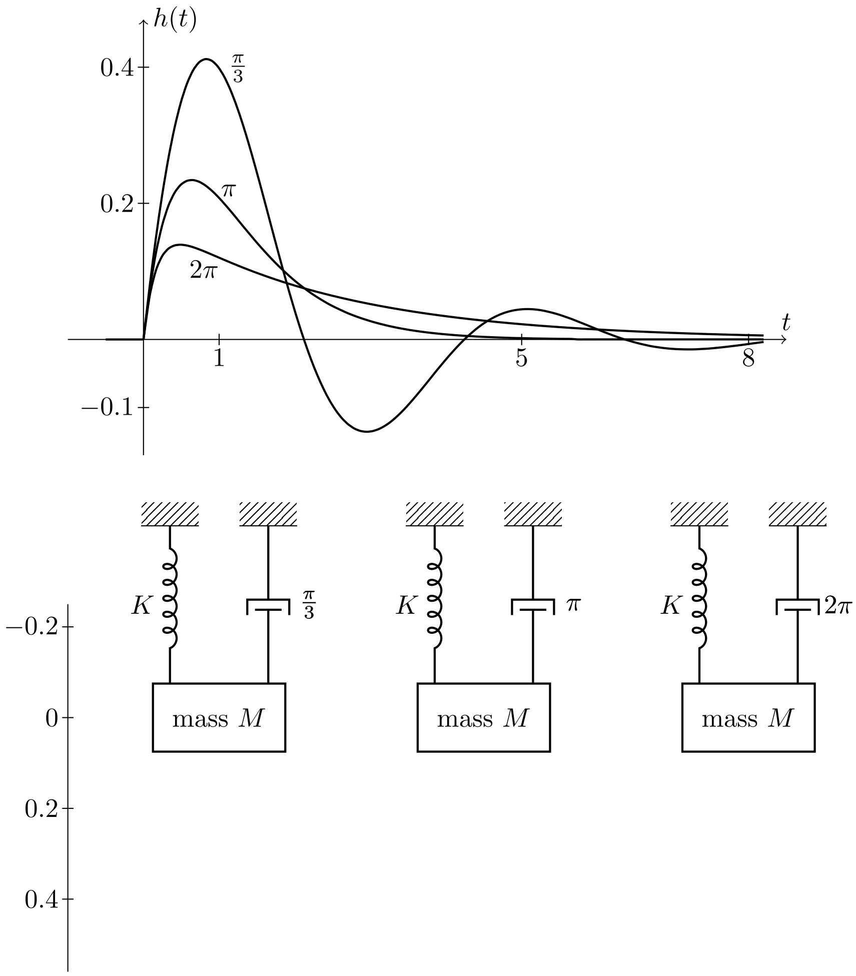

Figure 34: Impulse response of the mass spring damper with $M = 1$, $K = \frac{\pi^2}{4}$ and damping constant $B = \frac{\pi}{3}$ (underdamped), $B = \sqrt{4KM} = \pi$ (critically damped) and $B = 2\pi$ (overdamped).