Chapter 2

Signal Types and Operations

Firstly, what is a signal? It is the time evolution of a quantity. As examples, all of the following can be considered as signals

- The price of a share in a publically listed company;
- The level of water in a reservoir;
- The speed of a car;
- The temperature in a room;
- The voltage driving the speaker in a mobile telephone.

There are many other instances. The point is that for all the above cases it is easy to imagine engineering scenarios in which these signals are relevant.

For example, in the case of a car-speed signal (perhaps measured as a voltage from an electronic speedometer), an engineer might be required to design circuitry to process it, together with a 'commanded-speed' (set by the car driver), in order to synthesise a third signal that drives an actuator connected to the accelerator, and which makes the measured-speed signal identical to the commanded-speed signal. This would be a car 'cruise-control' system.

In order to address this (and other) types of engineering problems, certain mathematical tools are required. Therefore, a signal is most commonly (and abstractly) thought of as a mathematical function $f(t)$ of time t which can be any real number; see figure 2.1 for a graphical illustration of this.

Typically, the origin at time $t = 0$ is thought of as being 'now', with $f(t)$ for $t > 0$ being thought of as the future time evolution of the signal, and $f(t)$ for $t < 0$ being the past time evolution.

It should be pointed out that this description of a signal as a function $f(t)$ defined on the real number line (denoted by $\mathbf{R}$ in this book) is sometimes referred to as a 'continuous time signal' as opposed to a 'discrete time signal', which is only defined for $t \in \mathbf{Z}$, the set of integers. In this book, we will not meet such discrete time signals until chapter 6.

2.1 Types of Signals

Clearly, within this broad definition of a signal, there are many different possibilities (surely the time evolution of water in a large reservoir is qualitatively and quantitatively very different to the time evolution of a mobile telephone speaker voltage), and so it is useful to be able to classify signals and/or gain understanding by considering simplified but common signal types.

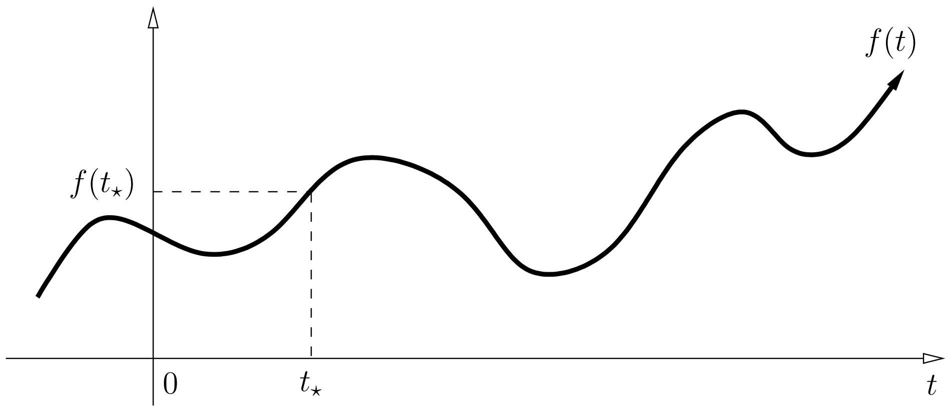

Figure 2.1: *Mathematical representation of a continuous time signal (i.e. the time evolution of a quantity) as a function $f(t)$ defined on the real line $\mathbf{R}$.*

2.1.1 Constant Signal

This is perhaps the simplest signal, which is just a constant value A, and therefore is defined as

$$f(t) \triangleq A \in \mathbf{R}, \qquad t \in (-\infty, \infty).$$

See figure 2.2 for an illustration of this signal.

Figure 2.2: *Constant Signal.*

2.1.2 Unit Step Signal

This signal, a natural extension of the previous one, is such that it is constant and zero up until 'now' (the origin), and then constant and non-zero in the future (the positive real axis). It is given the special function name $\mathbf{1}(t)$ and is defined as

$$\mathbf{1}(t) \triangleq \begin{cases} 1 & ; t \geq 0 \\ 0 & ; t < 0. \end{cases} \tag{2.1}$$

It is illustrated in figure 2.3.

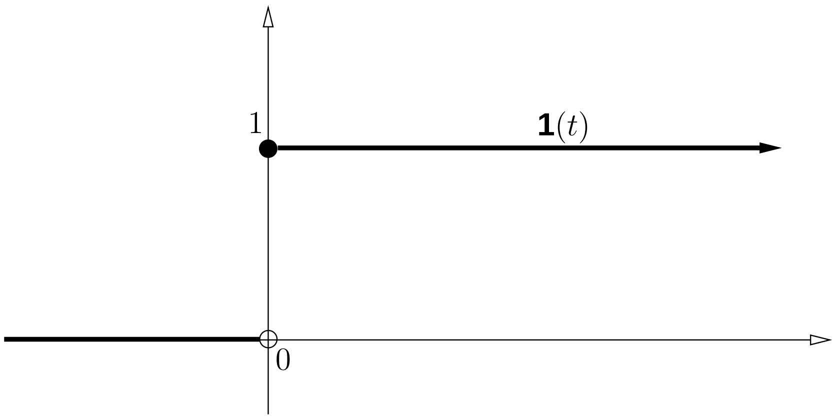

Figure 2.3: *Step Signal.*

2.1.3 Ramp Signal

This is a linearly increasing signal which is given the special symbol $r(t)$ and is defined as

$$r(t) \triangleq \begin{cases} t & ; t \geq 0 \\ 0 & ; t < 0. \end{cases}$$

It is illustrated in figure 2.4. This signal provides the first example of how one signal may be derived

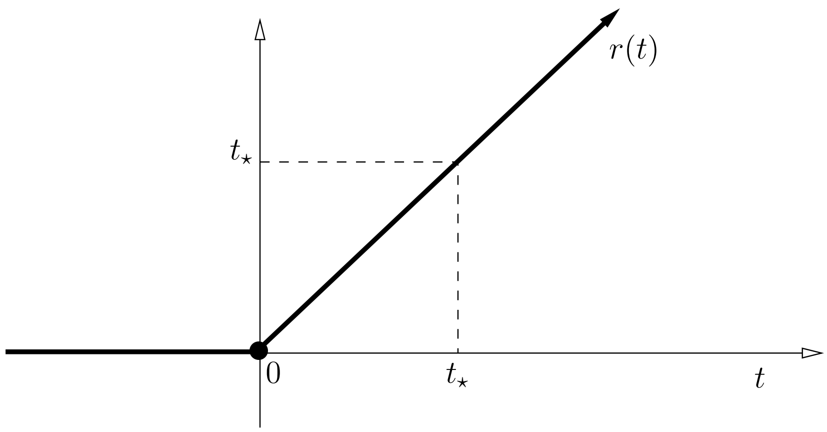

Figure 2.4: *Ramp Signal*

from another, since the ramp $r(t)$ is the cumulative area under the step signal $\mathbf{1}(t)$ up until time t. That is

$$r(t) = \int_{-\infty}^{t} \mathbf{1}(\sigma)\, \mathrm{d}\sigma. \tag{2.2}$$

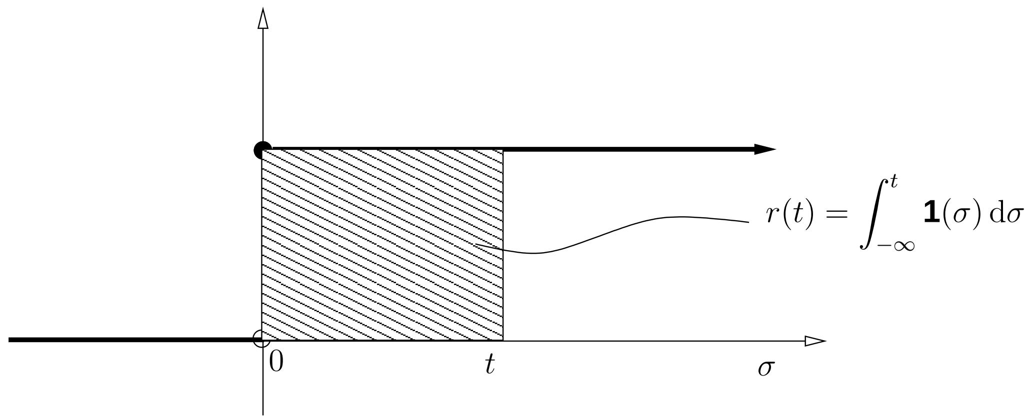

Figure 2.5: *Ramp signal r as the cumulative area under the step signal s.*

This principle is illustrated in figure 2.5. In fact, the relationship (2.2) highlights an important mathematical point which will arise many times in this book, and it involves the so-called 'dummy variable' σ which suddenly appeared in the integral (2.2). Why is it introduced as the argument (input variable) of the function $\mathbf{1}$ when previously the symbol t (representing time) was used?

The reason is that the symbol t is being employed as the argument for r in (2.2), and hence it must also be used as the upper limit of integration in (2.2). In this case, the symbol t cannot also be used as the 'moving' variable of integration within the integrand of (2.2). Another symbol has to be chosen, and in (2.2) σ was used as this new variable, but anything else (like x or μ, *but not* t) would have been suitable.

Also, as a consequence of (2.2) and the fundamental principle that differentiation is the inverse operation to integration, the process of differentiating both sides of (2.2) provides another relationship between $r(t)$ and $\mathbf{1}(t)$

$$\frac{\mathrm{d}}{\mathrm{d}t} r(t) = \mathbf{1}(t).$$

That is, the time rate of change of the ramp signal is the step signal. Finally, $r(t)$ and $\mathbf{1}(t)$ are also related by the equation

$$r(t) = t \cdot \mathbf{1}(t).$$

2.1.4 Pulse Signal

An important signal in relation to the analysis of (amongst many other things) digital communications systems is the pulse signal $\Pi_\tau(t)$ defined as

$$\Pi_\tau(t) \triangleq \begin{cases} 1 & ; |t| \leq \tau \\ 0 & ; |t| > \tau. \end{cases} \qquad (2.3)$$

and which is illustrated in figure 2.6 to be zero everywhere, except in the region $[-\tau, \tau]$ symmetric about the origin, where it takes the value one. Its significance stems from its ability to model the zero-to-one transitions inherent to the transmission of binary data.

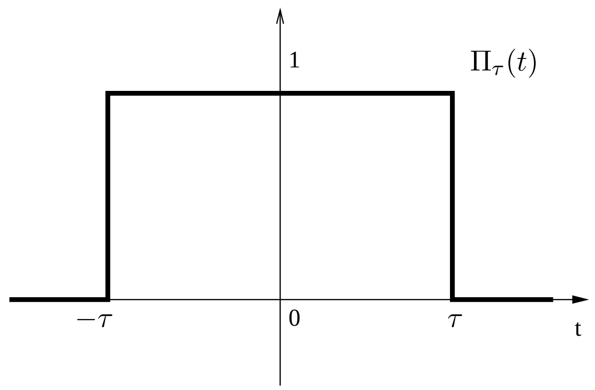

Figure 2.6: *Pulse Signal* $\Pi_\tau(t)$.

2.1.5 (Dirac) Delta Function

This signal, which is also known as the (unit) impulse signal, is the most difficult to characterise because, strictly speaking, it cannot be written as a bona-fide mathematical function. Why? Because the value of a (Dirac) delta signal, written as the special symbol $\delta(t)$, is not defined for the point $t = 0$. More specifically

$$\delta(t) \triangleq \begin{cases} 0 & ; t \neq 0 \\ \text{Undefined} & ; t = 0 \end{cases}$$

Although $\delta(t)$ is not defined at $t = 0$, it is useful to imagine that $\delta(0) = +\infty$ so that a graphical representation of $\delta(t)$ is as shown in figure 2.7; the upward arrow at $t = 0$ indicates the (qualitative) $+\infty$ value.

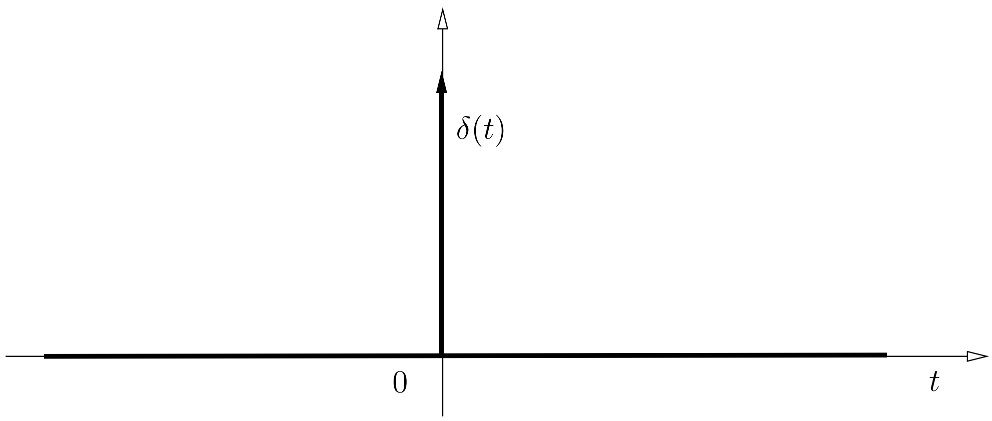

Figure 2.7: *(Dirac) Delta Signal* $\delta(t)$

The point of this signal is that it is often useful to imagine a scenario in which a 'unit' of a quantity can be delivered in zero time; ie. instantaneously. In fact, the 'Dirac' in the name refers to the Quantum Physics pioneer Paul Dirac who used this δ function to represent electrons as being units of charge occupying an infinitesimally small (i.e. zero) amount of space.

Note also that the Dirac epithet is used to discriminate it from another sort of δ function called the Kronecker delta, which is well defined at the origin, and which will be discussed later in relation to digital signal processing methods.

Moving on with the discussion of the Dirac delta, although it is not defined at $t = 0$ (and is therefore not, strictly speaking, a function), it is still possible to give a rigorous mathematical characterisation of $\delta(t)$. This is achieved by defining $\delta(t)$ not by its value for various choices of t, but by how it acts relative to other functions. Specifically, $\delta(t)$ may be characterised as a relationship which, for any continuous function[1] $f(t)$, allows the following integral relationship to hold

$$\int_{-\infty}^{\infty} f(t)\,\delta(t)\,\mathrm{d}t = f(0). \tag{2.4}$$

That is, $\delta(t)$ acts so as to 'pick out' the value of $f(t)$ at the point $t = 0$ (where we imagine that $\delta(t)$ becomes infinite) and so as to ignore values of $f(t)$ where $t \neq 0$ (and where $\delta(t) = 0$ so that $f(t)\,\delta(t) = 0$). For the interested reader, this means that $\delta(t)$ is technically known as a 'distribution' rather than as a function.

In fact, one of the key reasons for using the $\delta(t)$ in signal analysis problems is that it can greatly simplify calculations by virtue of the defining property (2.4). For example, suppose that $f(t)$ is so difficult that for some general signal $g(t)$ the integral

$$\int_{-\infty}^{\infty} f(t)g(t)\,\mathrm{d}t$$

is too complicated to evaluate or to gain insight into. In this case, a possible remedy is to consider the special case in which $g(t) = \delta(t)$ so that

$$\int_{-\infty}^{\infty} f(t)g(t)\,\mathrm{d}t = \int_{-\infty}^{\infty} f(t)\delta(t)\,\mathrm{d}t = f(0)$$

which is then easily evaluated.

To further elaborate on the sense behind the relationship (2.4), suppose that $\delta(t)$ is approximated by the function $K_\epsilon(t)$ defined as

$$\delta(t) \approx K_\epsilon(t) = \begin{cases} 1/\epsilon & ; |t| \le \epsilon/2 \\ \\ 0 & ; |t| > \epsilon/2 \end{cases}$$

and which is illustrated in figure 2.8. Then, since $K_\epsilon(t) = 0$ for $|t| > \epsilon/2$, then also $f(t)K_\epsilon(t) = 0$ for $|t| > \epsilon/2$ and hence

$$\begin{aligned} \int_{-\infty}^{\infty} f(t)\,\delta(t)\,\mathrm{d}\tau &\approx \int_{-\infty}^{\infty} f(t)K_\epsilon(t)\,\mathrm{d}t & (2.5) \\ &= \int_{-\epsilon/2}^{\epsilon/2} f(t)\frac{1}{\epsilon}\,\mathrm{d}t. & (2.6) \end{aligned}$$

Now, if ϵ is made very small, then since $f(t)$ is meant to be continuous, $f(t)$ should not vary by very much in the tiny interval $(-\epsilon/2, \epsilon/2)$, so that $f(t)$ is then approximately the constant value $f(t) \approx f(0)$ for all $t \in (-\epsilon/2, \epsilon/2)$. In this case

$$\int_{-\epsilon/2}^{\epsilon/2} f(t)\frac{1}{\epsilon}\,\mathrm{d}t \approx \frac{f(0)}{\epsilon}\int_{-\epsilon/2}^{\epsilon/2} \mathrm{d}t = \frac{f(0)}{\epsilon}\cdot\epsilon = f(0).$$

[1] See appendix 2.A.2 for revision on this topic.

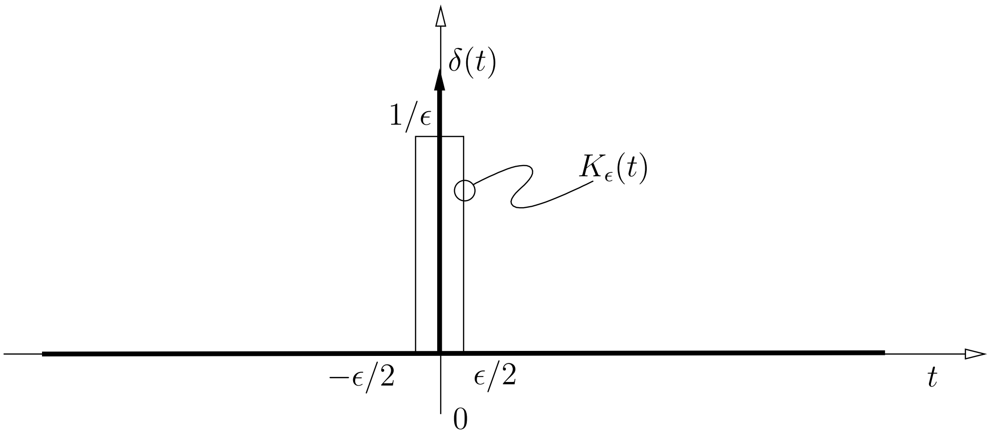

Figure 2.8: *Signal* $K_\epsilon(t)$ *that approximates* $\delta(t)$

Therefore, since the approximation $f(t) \approx f(0)$ for $t \in (-\epsilon/2, \epsilon/2)$ gets better as $\epsilon \to 0$ then

$$\lim_{\epsilon \to 0} \int_{-\epsilon/2}^{\epsilon/2} f(t) \frac{1}{\epsilon} \, \mathrm{d}t = f(0).$$

Consequently, if we imagine [2]

$$\delta(t) = \lim_{\epsilon \to 0} K_\epsilon(t)$$

then from (2.5)

$$\int_{-\infty}^{\infty} f(t)\delta(t) \, \mathrm{d}t = \int_{-\infty}^{\infty} f(t) \lim_{\epsilon \to 0} K_\epsilon(t) \, \mathrm{d}t = \lim_{\epsilon \to 0} \int_{-\infty}^{\infty} f(t) K_\epsilon(t) \, \mathrm{d}t = f(0).$$

This is also why $\delta(t)$ is also known as the *unit* impulse, since the area underneath it is unity:

$$\int_{-\infty}^{\infty} \delta(t) \, \mathrm{d}t = \lim_{\epsilon \to 0} \int_{-\infty}^{\infty} K_\epsilon(t) \, \mathrm{d}t = \lim_{\epsilon \to 0} \int_{-\epsilon/2}^{\epsilon/2} \frac{1}{\epsilon} \, \mathrm{d}t = \lim_{\epsilon \to 0} \frac{\epsilon}{\epsilon} = 1. \tag{2.7}$$

Notice also that apart from the justification according to $K_\epsilon(t)$, the relationship between the far left and far right hand sides of (2.7) is nothing more than the definition (2.4) with the substitution $f(t) = 1$.

Finally, also notice that there is an integral relationship between the unit step $\mathbf{1}(t)$ and the Dirac delta $\delta(t)$ which is given as

$$\mathbf{1}(t) = \int_{-\infty}^{t} \delta(\sigma) \, \mathrm{d}\sigma \tag{2.8}$$

so that, by differentiating both sides of the above equation, there is also a differential relationship between $\mathbf{1}(t)$ and $\delta(t)$ given as

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathbf{1}(t) = \delta(t).$$

[2] See appendix 2.A.1 for revision of what that $\lim$ operator means.

2.1.6 Periodic Signal

A periodic signal is one which consists of the same wave-shape repeated every T seconds and infinitely often. Mathematically then, a signal is periodic if

$$f(t) = f(t+T), \qquad \text{for all } t.$$

See figure 2.9 for an illustration of a particular example of a periodic signal.

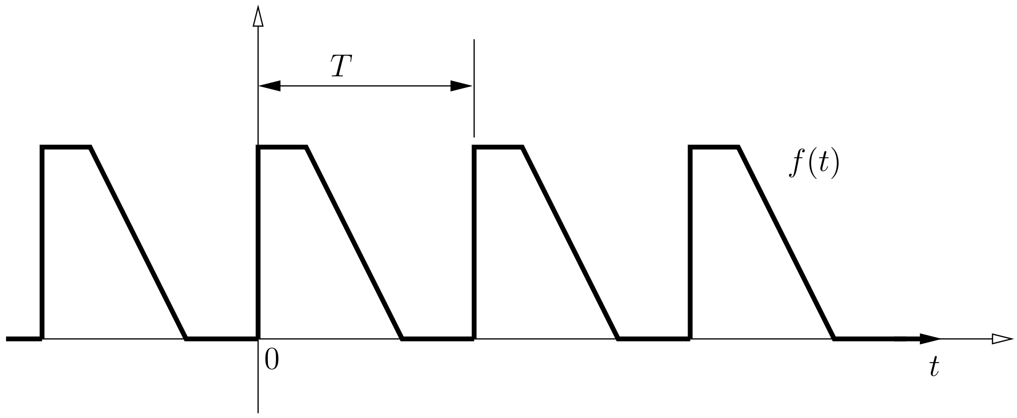

Figure 2.9: *A Periodic Signal.*

Example 2.1 The signal

$$f(t) = \sin \omega t$$

is periodic with period $T = 2\pi/\omega$ since by basic rules of trigonometry

$$\begin{aligned}
f(t+T) &= \sin \omega \left(t + \frac{2\pi}{\omega} \right) \\
&= \sin(\omega t + 2\pi) \\
&= \sin \omega t \cdot \underbrace{\cos 2\pi}_{=1} + \underbrace{\sin 2\pi}_{=0} \cdot \cos \omega t \\
&= \sin \omega t = f(t).
\end{aligned}$$

■

An important point is that any (possibly weighted) sum of periodic signals, all with the same period T, is also periodic of period T. That is, if $f(t) = f(t+T)$ and $g(t) = g(t+T)$ then for any $\alpha, \beta \in \mathbf{R}$

$$h(t) = \alpha f(t) + \beta g(t) = \alpha f(t+T) + \beta g(t+T) = h(t+T).$$

2.1.7 Even and Odd Signals

An even signal is one which is symmetric about the y axis (vertical axis through the origin) when plotted. That is, $f(t)$ is an even signal if

$$f(t) = f(-t), \qquad \text{for all } t.$$

An odd signal, is one which is symmetric about the origin. That is, $g(t)$ is an odd signal if

$$g(t) = -g(-t), \qquad \text{for all t.}$$

See figure 2.10 for illustrations of particular examples of even and odd signals.

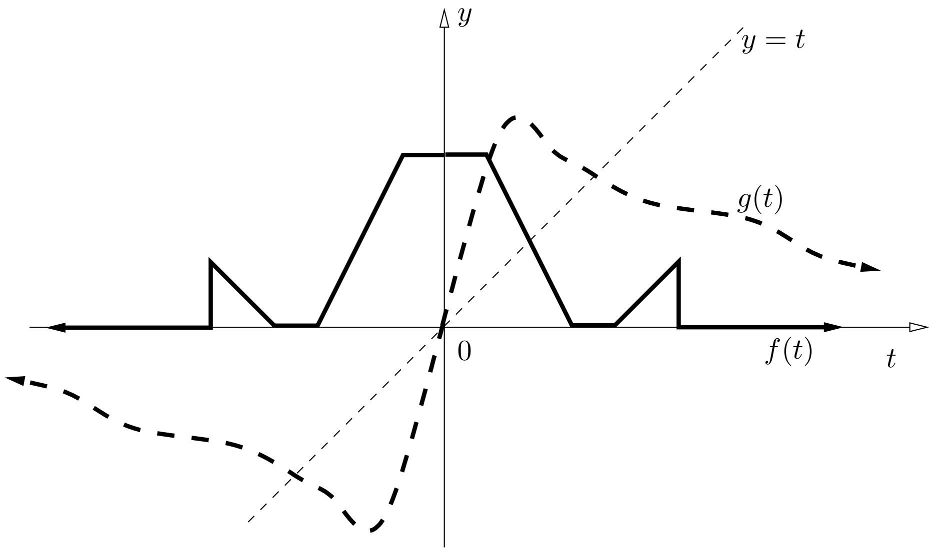

Figure 2.10: *An even signal $f(t)$ and an odd signal $g(t)$.*

Example 2.2 The signal $f(t) = \cos \omega t$ is even while the signal $g(t) = \sin \omega t$ is odd.

■

2.1.8 Exponential Signals

It so happens that the signals arising from a broad class of engineering systems contain components which are exponential. That is, they contain a signal of the form

$$f(t) = \mathbf{1}(t) \cdot Ae^{\alpha t} = \begin{cases} Ae^{\alpha t} & ; t \geq 0 \\ 0 & ; t < 0 \end{cases}$$

where α is a real valued constant. If $\alpha > 0$ then $e^{\alpha t}$ is exponentially increasing, while if $\alpha < 0$ then $e^{\alpha t}$ is exponentially decreasing. See figure 2.11 for an illustration of this. Note in particular that when $\alpha = 0$ then $e^{\alpha t} = e^{0.t} = e^0 = 1$ so the exponential signal becomes the step signal $\mathbf{1}(t)$.

2.1.9 Exponential Signals: Complex Representations

As well as containing exponential components, signals from physical systems also often contain oscillating components, so that they are of the form

$$f(t) = \mathbf{1}(t) \cdot Ae^{\alpha t} \cos(\omega t + \phi) = \begin{cases} A\, e^{\alpha t} \cos(\omega t + \phi) & ; t \geq 0 \\ 0 & ; t < 0 \end{cases}$$

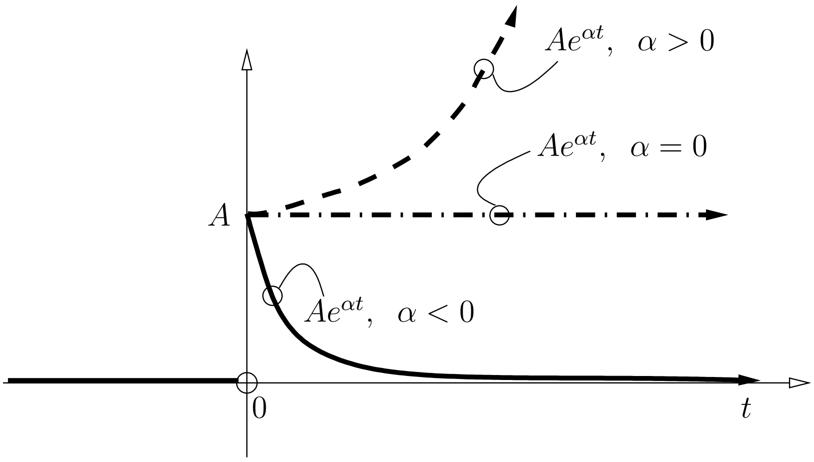

Figure 2.11: *Exponential Signal $e^{\alpha t}$ for various choices of α.*

This is an oscillating signal, with oscillation period $T = 2\pi/\omega$ seconds, exponential envelope $A\,e^{\alpha t}$ and co-sinusoidal shape with phase offset ϕ radians. See figure 2.12 for an illustration of this. An alternative and very useful way of representing a signal $f(t)$ of this sort arises by using the facts that for any numbers A and B

$$e^{A+B} = e^A \cdot e^B$$

and also if j represents $\sqrt{-1}$ (the complex square root of -1) then

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

by De–Moivre's Theorem

$$Ae^{\alpha t + j(\omega t+\phi)} = A\,e^{\alpha t}\cdot e^{j(\omega t+\phi)} = Ae^{\alpha t}\left[\cos(\omega t+\phi) + j\sin(\omega t+\phi)\right]$$

and hence, for $t \geq 0$

$$f(t) = Ae^{\alpha t}\cos(\omega t+\phi) = \text{Real}\left\{Ae^{\alpha + j(\omega t+\phi)}\right\}.$$

Now, it so happens that one of the most fundamental principles in the area of signals and systems is that in many cases, the frequency ω is unchanged by any physical effects, while the magnitude A, the exponential rate of decay α and the phase offset ϕ might all be altered. In this case, the book-keeping of A, α and ϕ may be achieved by simply ignoring the ωt term above and just associating

$$f \Leftrightarrow Ae^{\alpha + j\phi}.$$

There are now two important cases. The first is one in which it is known that the exponential damping constant $\alpha = 0$, so that the amplitude of $f(t)$ is constant; see figure 2.13. In this case, the signal has a complex valued representation $f \Leftrightarrow Ae^{j\phi}$. That is, the single complex number $Ae^{j\phi}$ can be used as a shorthand representation for the signal $f(t)$ on the understanding that implies

$$f(t) = \text{Real}\left\{Ae^{j(\omega t+\phi)}\right\} = A\cos(\omega t+\phi).$$

This technique is often called the 'phasor' representation of a sinusoidally shaped signal.

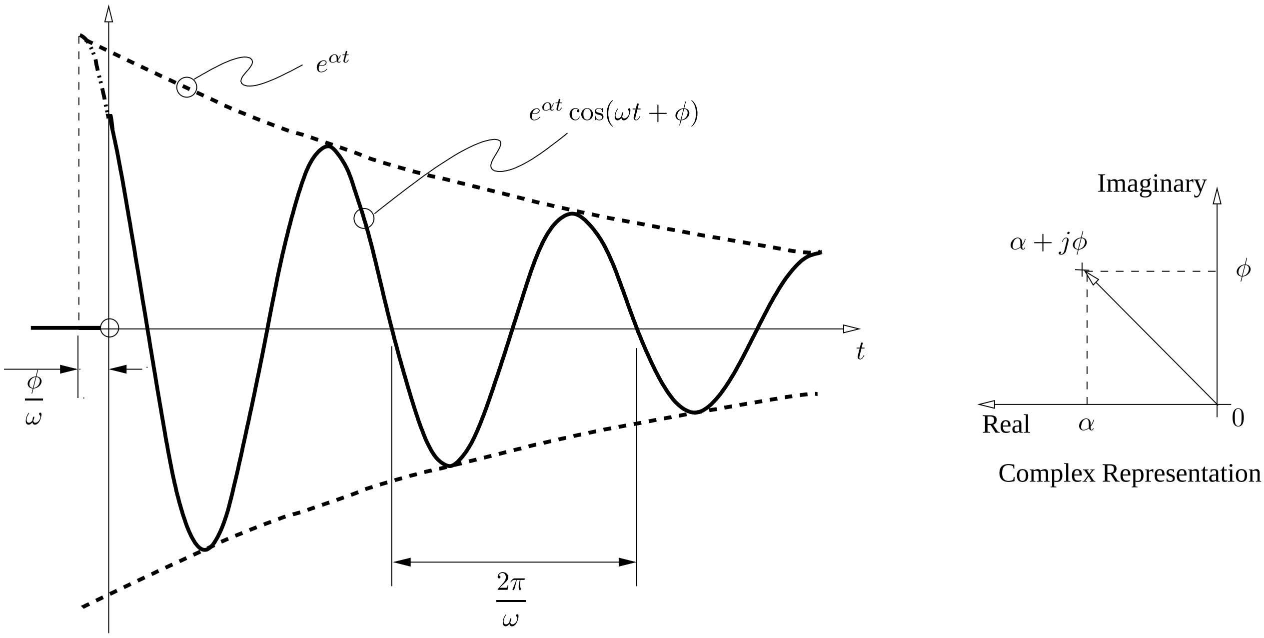

Figure 2.12: *Generalised Exponential Signal Re*$\{e^{(\alpha+j\omega)t}\} = e^{\alpha t}\cos(\omega t)$ *for* $\alpha < 0$.

The other important scenario is when the amplitude has an exponential envelope $e^{\alpha t}$ that needs to be specified. In this situation, a complex representation $f \Leftrightarrow \alpha + j\phi$ can be used as shorthand for

$$f(t) = e^{\alpha t}\cos(\omega t + \phi) = \text{Real}\left\{e^{\alpha + j(\omega t + \phi)}\right\}.$$

This situation for the case of $\alpha < 0$ has already been shown in figure 2.12, and for the case of $\alpha > 0$ is shown in 2.14.

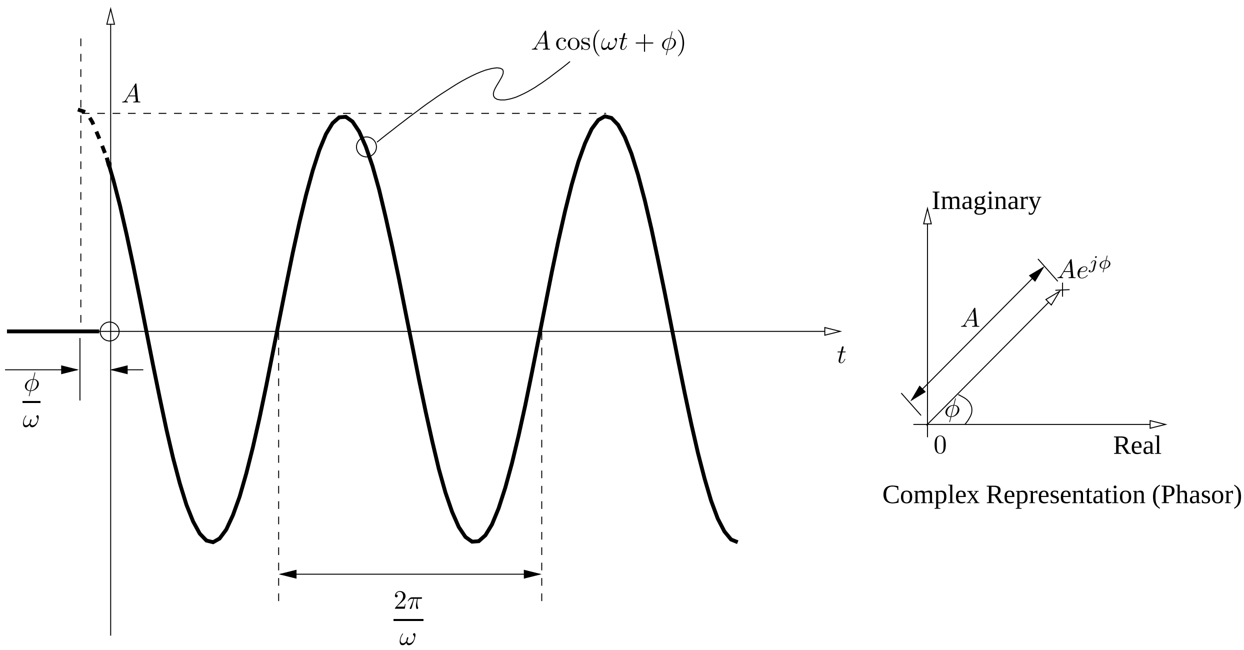

Figure 2.13: *Exponential Signal Re*$\{Ae^{j\omega t}\} = A\cos(\omega t + \phi)$.

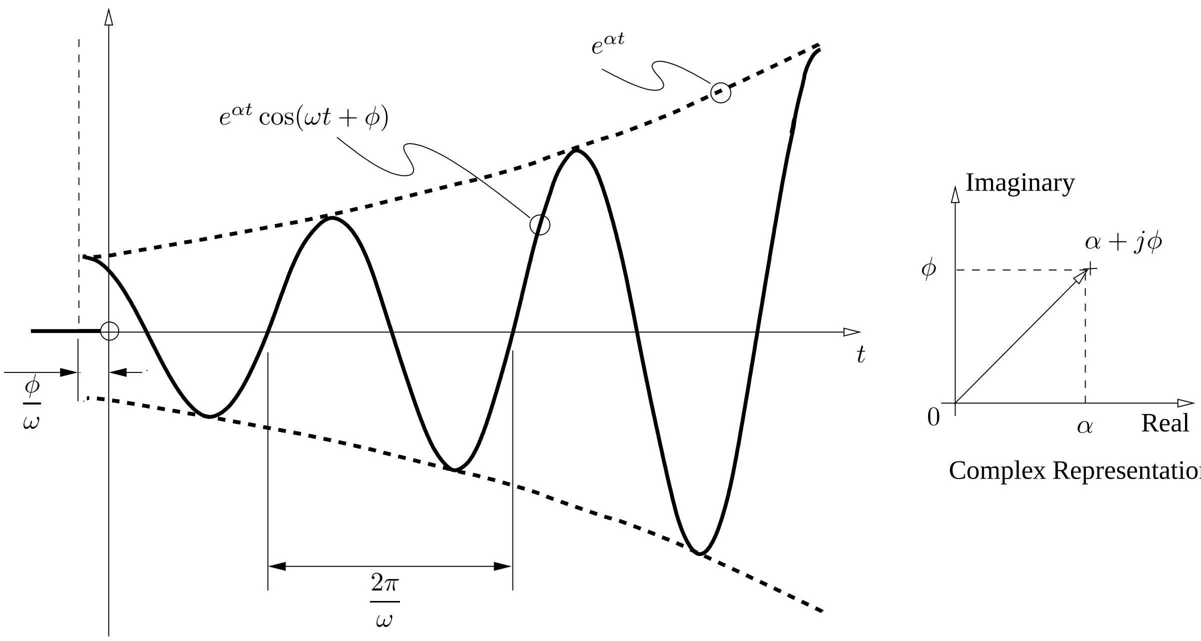

Figure 2.14: *Generalised Exponential Signal Re*$\{e^{(\alpha+j\omega)t}\} = e^{\alpha t}\cos(\omega t)$ *for* $\alpha > 0$.

2.2 Operations on Signals

The preceding section 2.1 introduced several basic 'building block' signals, from which more complicated signals can be derived via various fundamental operations which are defined in this section.

2.2.1 Magnitude Scaling

Probably the simplest signal operation is that of magnitude scaling, which is nothing more that multiplication by a constant $K \in \mathbf{R}$. That is, given an original signal $f(t)$, a new magnitude scaled one $g(t)$ is formed as

$$g(t) = Kf(t).$$

Depending on whether $|K| > 1$ or $|K| < 1$ this transforms the 'size' of a signal (respectively) up or down. Furthermore, according to whether $K < 0$ or $K > 0$ the signal is (again, respectively) flipped or not around the horizontal (time) axis. See figure 2.15 for an illustration of these operations.

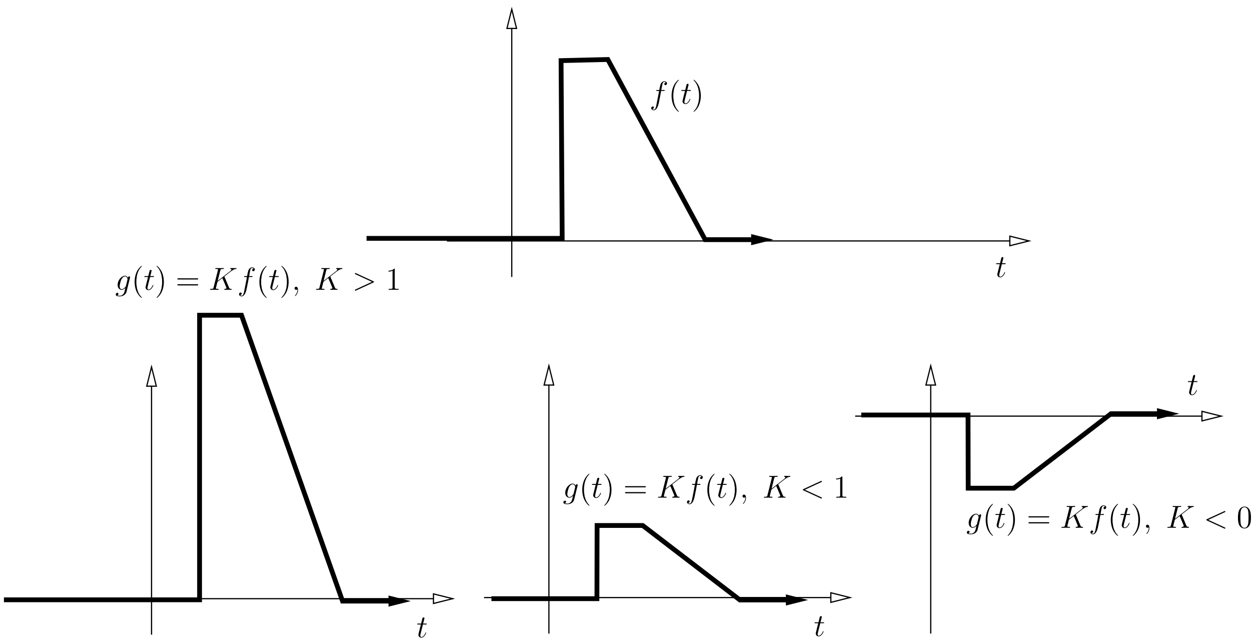

Figure 2.15: *The signal $g(t) = Kf(t)$ which is a magnitude scaled version of $f(t)$.*

2.2.2 Time Shifting (Translation)

Suppose that, given a signal $f(t)$, a new one $g(t)$ is formed from it as

$$g(t) = f(t - T).$$

Then $g(t)$ is simply $f(t)$ shifted to the right by T seconds; see figure 2.16 for an illustration of this. Another way of thinking about this operation is that $g(t) = f(t - T)$ is $f(t)$ *delayed* by T seconds, since (for example) the value $f(0)$ of the signal $f(t)$ at $t = 0$ is not taken on by $g(t)$ until time $t = T$, at which point $g(T) = f(T - T) = f(0)$.

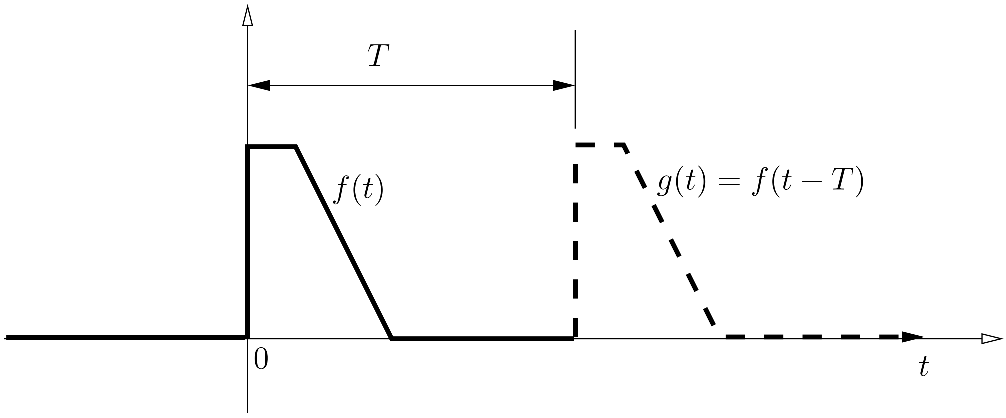

Figure 2.16: *The signal $g(t) = f(t-T)$ which is $f(t)$ shifted to the right (delayed) by T seconds.*

Notice also that if T is chosen as a negative number, then $g(t) = f(t-T)$ is $f(t)$ shifted to the left, or advanced in time; ie. events happen earlier in $g(t)$ than in $f(t)$.

This operation of time shifting is of particular interest with regard to its effect on Dirac delta signals. Firstly, note that according to the preceding discussion, the signal

$$f(t) = \delta(t-T)$$

should be a signal which is $\delta(t)$ shifted to the right by T seconds, and hence $\delta(t-T)$ should obey

$$\delta(t-T) = \begin{cases} 0 & ; t \neq T \\ \text{Undefined} & ; t = T. \end{cases}$$

See figure 2.17 for an illustration of this, which clarifies that intuitively it is reasonable to think of

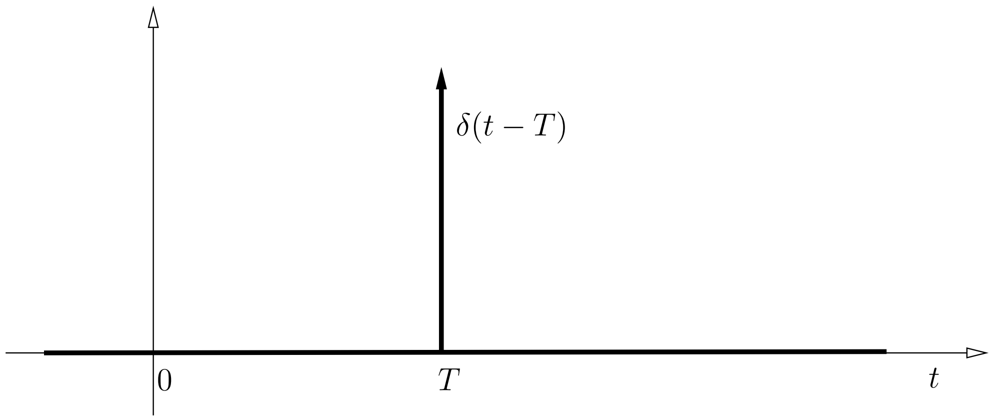

Figure 2.17: *The signal $f(t) = \delta(t-T)$.*

$\delta(t-T)$ as having all its 'mass' concentrated at $t = T$. In this case, according to the discussion of section 2.1.5, it is also reasonable to actually define $\delta(t-T)$ according to its action on a continuous function $f(t)$ via

$$\int_{-\infty}^{\infty} f(t)\delta(t-T)\,\mathrm{d}t = f(T). \tag{2.9}$$

That is, $\delta(t-T)$ is such that the integral (2.9) ignores all values of $f(t)$ for which $t \neq T$ and hence $\delta(t-T) = 0$ (so that $f(t)\delta(t-T) = 0$) but weights the value $f(t)$ at $t = T$ by the very large quantity $\delta(T)$ in such a way that the integral (2.9) just 'picks out' this value. This principle will turn out to be of great importance in later developments.

2.2.3 Time Reversal (Flipping)

For a general signal $f(t)$, this operation produces a new signal $g(t)$ by reversing the time direction

$$g(t) = f(-t)$$

and hence $g(t)$ is $f(t)$ flipped symmetrically about the y (vertical) axis through the origin; see figure 2.18 for an illustration of this operation. Of particular interest (especially when considering

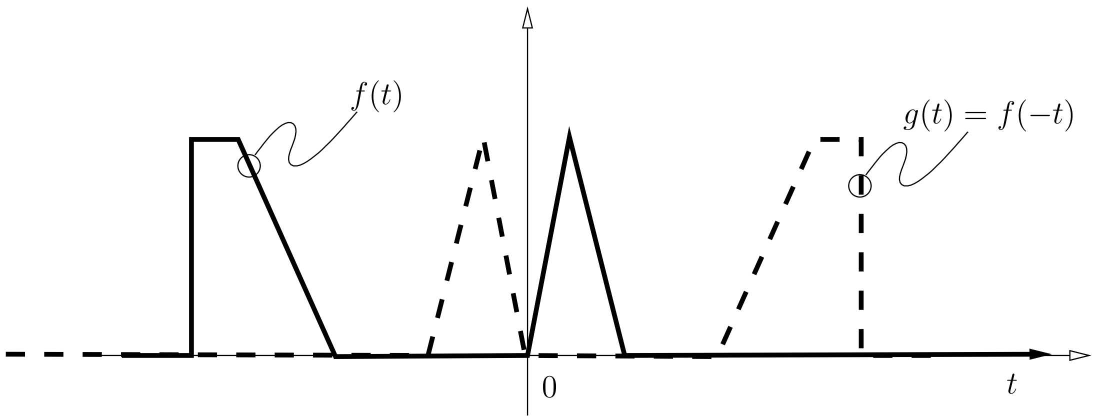

Figure 2.18: *Illustration of the operation of signal time reversal (flipping around the y axis).*

a fundamental principal called 'convolution' which will be discussed later) is the situation where this time reversal operation is combined with the time shifting operation. For example, the signal $g(t) = f(T-t)$ is $f(t)$ flipped and shifted to the right by T seconds; see figure 2.19. Note that, at first, the result shown in figure 2.19 might seem confusing since

1. By the discussion in section 2.2.2, $f(t-T)$ is $f(t)$ shifted to the *right* by T seconds, so $g(t) = f(T-t) = f(-(t-T))$ should be $f(t-T)$ mirrored about the y axis, and hence appearing to the *left* of the origin, whereas figure 2.19 shows it appearing moved to the right;

2. As another argument against shifting to the right, $g(t) = f(-t)$ is $f(t)$ flipped, and then $g(t+T) = f(T-t)$ should be $g(t)$ shifted to the left, which is again opposite to what is shown in figure 2.19.

Of course, the above arguments 1 and 2 are both flawed. Specifically, with argument 1, although the first step

$$g(t) = f(T-t)$$

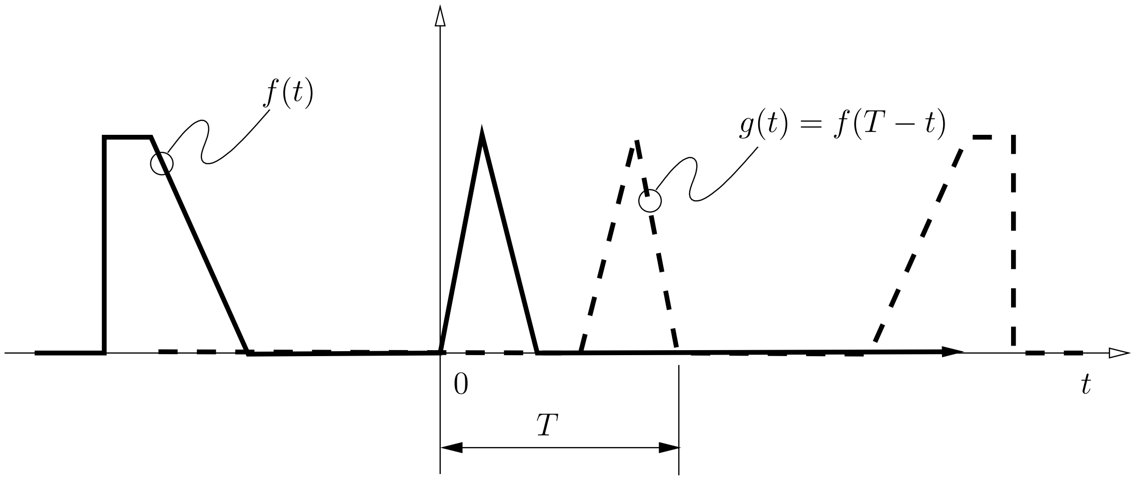

Figure 2.19: *Illustration of the operation of combined signal time reversal and time shifting.*

does indeed give a signal shifted to the right, careful thought shows that

$$g(-t) \neq f(-(T-t)).$$

and hence $f(t-T)$ is not just $g(t)$ flipped. In fact

$$g(-t) = f(T-(-t)) = f(T+t).$$

Similarly, with regard to argument 2 in which $g(t) = f(-t)$ is taken as a first step, it is in fact the case that

$$g(t+T) \neq f(-t+T)$$

as was intimated in argument 2. Actually,

$$g(t+T) = f(-(t+T)) = f(-t-T).$$

The point of raising these issues is that they are the source of common mistakes. It is important to be careful about what is the actual argument of a function.

2.2.4 Time Scaling

This operation is one in which the time axis of a signal is stretched or compressed by a factor $\alpha \in \mathbf{R}^+$. That is, a signal $g(t)$ is the α time-scaled version of a signal $f(t)$ if

$$g(t) = f(\alpha t).$$

In this case, by definition an event in the evolution of g occurring at time t occurs at time αt of f. Therefore, if $\alpha > 1$ then this corresponds to time compression since events happen earlier in the new signal g than they did in the original signal f. Vice-versa, if $\alpha < 1$ then time expansion takes place since αt is earlier (smaller) than t so that events are occurring later in g than they did in f; see figure 2.20 for an illustration of these principles.

Again, the Dirac delta function requires special attention. Since all the non-zero part of $\delta(t)$ occurs over zero width (at the origin), then what is the effect of scaling this zero width?

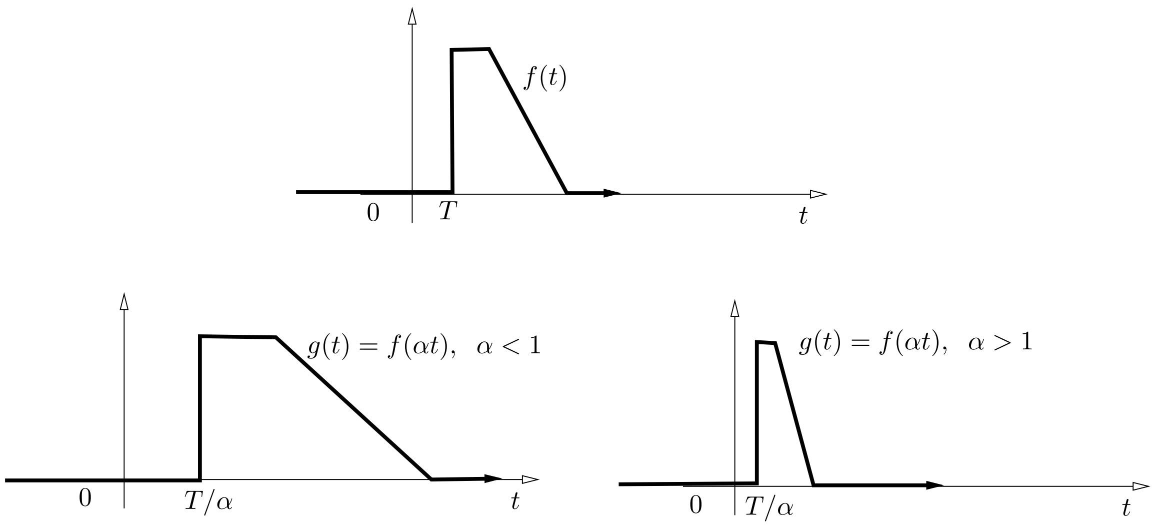

Figure 2.20: *Illustration of the time scaling operation for scale factor* $\alpha < 1$ *and scale factor* $\alpha > 1$.

The answer is found by appealing to the defining relationship (2.4). Specifically, using the change of variable $\sigma = \alpha t$ which implies that $\mathrm{d}\sigma = \alpha \mathrm{d}t$, then by (2.4)

$$\int_{-\infty}^{\infty} f(t)\delta(\alpha t)\,\mathrm{d}t = \int_{-\infty}^{\infty} f\left(\frac{\sigma}{\alpha}\right)\delta(\sigma)\frac{1}{\alpha}\,\mathrm{d}\sigma = \frac{1}{\alpha}f(0)$$

so that

$$\delta(\alpha t) = \frac{1}{\alpha}\delta(t). \tag{2.10}$$

That is, the time scaled delta function $\delta(\alpha t)$ is in fact another delta function, but this time with area $1/\alpha$. This makes intuitive sense since, as has already been explained, $\alpha > 1$ compresses the time axis, and hence the 'width' of the δ function, so this should lead to a decrease in area under the delta function by a factor $1/\alpha$.

2.3 Concluding Summary

We have now been introduced to some of the most fundamental ideas in the field of 'Signals and Systems'. Namely,

1. A signal is the time evolution of a quantity;
2. It can be represented as a mathematical function $f(t) : \mathbf{R} \to \mathbf{R}$;
3. There are some fundamental signal types,
   (a) Constant: $f(t) = K$;
   (b) Unit step: $f(t) = \mathbf{1}(t)$;
   (c) Unit ramp: $f(t) = r(t) = t \cdot \mathbf{1}(t)$;
   (d) Dirac Delta: $f(t) = \delta(t)$;

(e) Periodic of period T: $f(t) = f(t+T)$;

(f) Even $f(t) = f(-t)$ and odd $f(t) = -f(-t)$;

(g) Exponential: $f(t) = Ae^{\alpha t}$;

(h) Generalised Exponential: $f(t) = Ae^{\alpha t}\cos(\omega t + \phi)$.

4. There are some fundamental signal operations,

   (a) Magnitude Scaling: $g(t) = Kf(t)$;

   (b) Time shifting (translation): $g(t) = f(t - T)$;

   (c) Time reversal (flipping): $g(t) = f(-t)$;

   (d) Time scaling: $g(t) = f(\alpha t)$.

With all these ideas in hand, we will now progress in the following chapter to how we may represent, analyse and understand the relationship between two signals $f(t)$ and $g(t)$ that are governed the nature of some physical system.

2.4 Exercises

1. (a) Sketch $p(t) = \mathbf{1}(t+1) - \mathbf{1}(t-1)$
   (b) Show how the signals in figure 2.21 could be generated from $p(t)$
2. Show how the signals in figure 2.22 could be generated from $p(t) = \mathbf{1}(t+1) - \mathbf{1}(t-1)$ and $\mathbf{1}(t)$.
3. Sketch the following signals

   (a)
$$x(t) = \mathbf{1}(t+1) - 2\mathbf{1}(t-1) + \mathbf{1}(t-3).$$

   (b)
$$x(t) = (t+1)\mathbf{1}(t-1) - t\mathbf{1}(t) - \mathbf{1}(t-2).$$

   (c)
$$x(t) = e^{-t}\mathbf{1}(t) + e^{-t}\left[e^{2t-4} - 1\right]\mathbf{1}(t-2) - e^{t-4}\mathbf{1}(t-4).$$

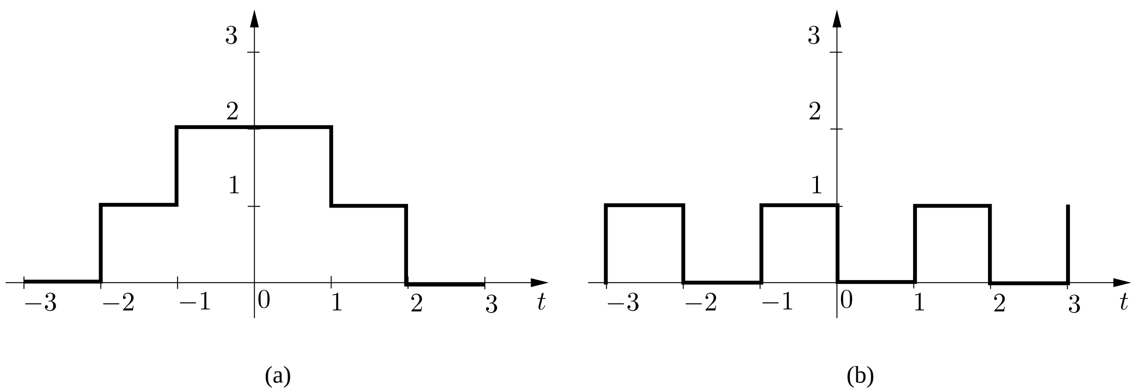

Figure 2.21: *Signals pertaining to question 1*

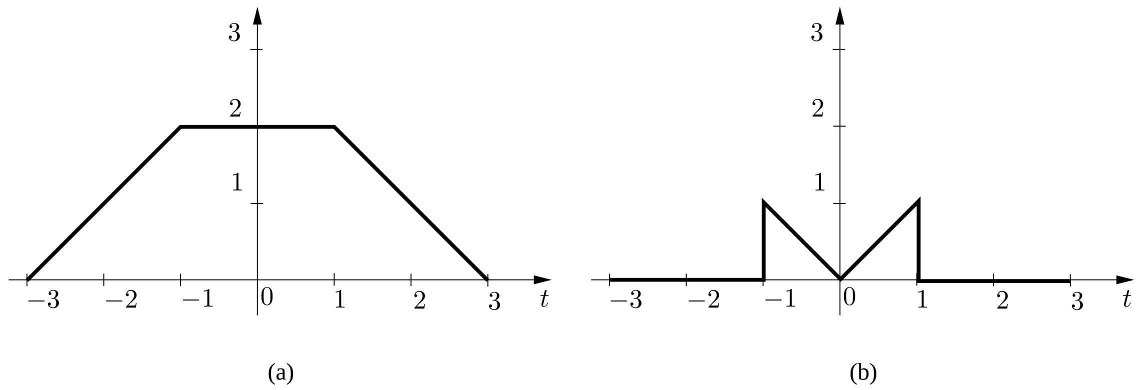

Figure 2.22: *Signals pertaining to question 2*

2.A Some Mathematical Background

2.A.1 The limit of a function

The limiting value $f_\star$ that a function $f(t)$ tends towards as t approaches a fixed value $t_\star$ is written as

$$f_\star = \lim_{t\to t_\star} f(t). \tag{2.11}$$

Notice that this limit $f_\star$ is not necessarily equal to $f(t_\star)$. For example, (see figure 2.3) with the step function $\mathbf{1}(t)$, if t approaches $t_\star = 0$ from above, then intuitively it would appear that $\lim_{t\to 0} = 1$ whereas $\mathbf{1}(0) = 0$. This sort of intuitive idea actually needs to be made more precise by a formal definition of what is meant by a limit, and this is achieved as follows.

The limit (2.11) exists if, for any $\epsilon > 0$, then no matter how small this ϵ may be, there exists a corresponding $\gamma > 0$ such that if

$$|t - t_\star| < \gamma$$

then

$$|f(t) - f_\star| < \epsilon.$$

That is, if the limit (2.11) exists, then I can make $f(t)$ as close as I like to $f_\star$ by making t close enought to $t_\star$.

Note that according to this definition

$$\lim_{t\to 0} \mathbf{1}(t)$$

does not exist! For any tiny region $t \in (-\gamma, \gamma)$ around 0, the function $\mathbf{1}(t)$ takes on values in the full range $[0, 1]$, regardless of how small γ is, and hence there is no number $\mathbf{1}_\star$ for which $|\mathbf{1}(t) - \mathbf{1}_\star|$ can be made to be less than an arbitrarily small ϵ.

Furthermore, it is also not necessarily true that $\lim_{t\to t_\star}$ is equal to $f(t_\star)$. For example, referring to the Dirac δ figure 2.7 illustrates that

$$\lim_{t\to 0} \delta(t) = 0.$$

However, $\delta(0)$ is undefined, although we think of it as being $+\infty$.

2.A.2 Continuous Functions

Continuous functions are the ones for which

$$\lim_{t\to t_\star} f(t) = f(t_\star)$$

and for any value of $t_\star$. They are called continuous, because for the above to be true, then $f(t)$ must not contain any abrupt ‘jumps’. For example then, neither $\mathbf{1}(t)$ nor $\delta(t)$ are continuous since, in the first case $\lim_{t\to 0} \mathbf{1}(t)$ does not exist, and in the second case $\lim_{t\to 0} \delta(t) = 0 \neq \delta(0)$. On the other hand, the ramp function $r(t)$ is continuous since (see figure 2.4)

$$\lim_{t\to t_\star} r(t) = \begin{cases} 0 = r(t_\star) & t \le 0 \\ \\ t_\star = r(t_\star) & t > 0. \end{cases}$$