Convolution! (CDT-14)

Preprint · October 2019

DOI: 10.13140/RG.2.2.27310.25925

Luciano da F. Costa
University of São Paulo

Convolution!
(CDT-14)

Luciano da Fontoura Costa
luciano@ifsc.usp.br

São Carlos Institute of Physics – DFCM/USP

October 10, 2020

Abstract

The convolution between two functions, yielding a third function, is a particularly important concept in several areas including physics, engineering, statistics, and mathematics, to name but a few. Yet, it is not often so easy to be conceptually understood, as a consequence of its seemingly intricate definition. In this text, we develop a conceptual framework aimed at hopefully providing a more complete and integrated conceptual understanding of this important operation. In particular, we adopt an alternative graphical interpretation in the time domain, allowing the shift implied in the convolution to proceed over free variable instead of the additive inverse of this variable. In addition, we discuss two possible conceptual interpretations of the convolution of two functions as: (i) the '*blending*' of these functions, and (ii) as a quantification of '*matching*' between those functions. We also discuss the convolution in terms of the frequency domain, through the convolution theorem, with emphasis on the particularly important case of convolution between two sinusoidal functions. The potential of application of these concepts and methods is illustrated with respect to dynamic systems and signal filtering.

'L'onda mai è sola, ma è mista di tant'altre onde...'

Leonardo da Vinci.

1 Introduction

Convolution... or, should we shout, *convolution!* This seemingly intricate operation is, at the same time, surprisingly important in physics, engineering, statistics, mathematics, and many other theoretical and applied areas.

Given two functions, convolution combines them in a specific (linear and equitable) way, facilitating several key mathematical operations, ranging from integro-differential operations to pattern recognition – passing through interpolation, system modeling, and filtering – to name but a few more immediate applications.

It is important to notice convolution is not a function, but a *binary operation* in the sense that it receives two functions as input, producing a third function as output, as illustrated in Figure 1.

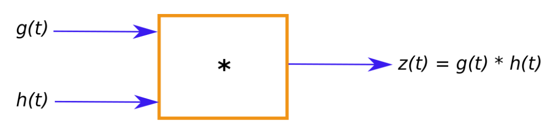

Figure 1: The *convolution* between two real/complex functions $g(t)$ and $h(t)$ is a third real/complex function $w(t)$ that can often be understood as a combination of the two input functions. Convolution is commutative, so there is no need to distinguish between its two inputs.

The Dirac's delta 'function' acts as the *identity element*, in the sense that, when convolved with another function $g(t)$, yields that same function. However, there is no guaranteed inverse function of $g(t)$, so that the binary operation of convolution does not formally constitute a *group* (e.g. [1]). Yet, convolution is a *bilinear operation*, in the sense that when one of the functions is kept fixed, the convolution acts linearly on the other function, and vice-versa.

Despite its importance, convolution can seem intricate at times. This probably stems from its own definition, in particular the variable transformation $\tau - t$, the fact that convolution corresponds to an infinite integral, the involvement of two functions, and the fact that complex functions are often adopted. Interestingly, even if only as a coincidence, the own related term *convoluted* is traditionally taken to be a synonymous of intricate.

Better understanding and application of mathematical concepts and methods can greatly benefit from respective intuitive and conceptual assimilations. Fortunately, there are several intuitive interpretations that can help one to develop a more conceptual and complete understanding of the type of operations performed by the convolution between two functions. In this work, we resource to several approaches so as to contribute to that finality. More specifically, we start by presenting an alternative, but equivalent, definition of convolution that allow the variable shifting required by this operation to be performed without involving the additive inverse of that variable, as typically implied by the more traditional definition of convolution. In addition, we also discuss convolution as the 'blending' or 'mixing' between the two involved functions, and also as 'matching' between those functions. The convolution is also approached considering the frequency domain, as allowed by the convolution theorem and the Fourier transform (which are also briefly presented). This approach allows one to realize that the convolution between two sinusoidal functions if null except for the case in which the two function have *exactly* the same frequency.

The consideration of several properties and applications of this important binary operation, including linear systems modeling and filtering, can provide further insights and consolidation of a more integrated and complete conceptual understanding of convolution.

This constitute the main objective of the current text.

2 Convolution

The *convolution* between two functions $g(t)$ and $h(t)$, both mapping from real values t into complex images, is traditionally defined as

$$[g * h](\tau) = \int_{-\infty}^{\infty} g(t)h(\tau - t)dt \tag{1}$$

If we make the variable substitution $u = \tau - t$, we have $du = -dt$, from which follows that

$$[g * h](\tau) = \int_{-\infty}^{\infty} g(\tau - u)h(u)du \tag{2}$$

implying that convolution is *commutative*, i.e. $g(t) * h(t) = h(t) * g(t)$.

Though this definition has been frequently considered, it involves a combination of the value τ with $-t$, which is perhaps not very intuitive, as the addition of τ in the argument of $h()$, i.e. $h(\tau - t)$, implies this function being displaced to the right, and not to the left as it would happen otherwise if t were not negated, i.e. $h(t + \tau)$.

A more intuitive expression for the convolution can be obtained by making the variable substitution $u = -t$, so that

$$\boxed{g(t) * h(t) = [g * h](\tau) = \int_{-\infty}^{\infty} g(-t)h(t + \tau)dt} \tag{3}$$

Now, we have that the variable τ, which is external to the integral, is added to t instead of $-t$. This is achieved at the expense of reflecting $g(t)$ into $g(-t)$, but this is more intuitive since this reflection does not involve shiftings by τ.

The above alternative definition of convolution, henceforth adopted in this work, also allows a simple graphical explanation of the operation implemented by the convolution between the two original functions $g(t)$ and $h(t)$, which is illustrated in Figure 2 for an even $g(t)$ (i.e. $g(t) = g(-t)$).

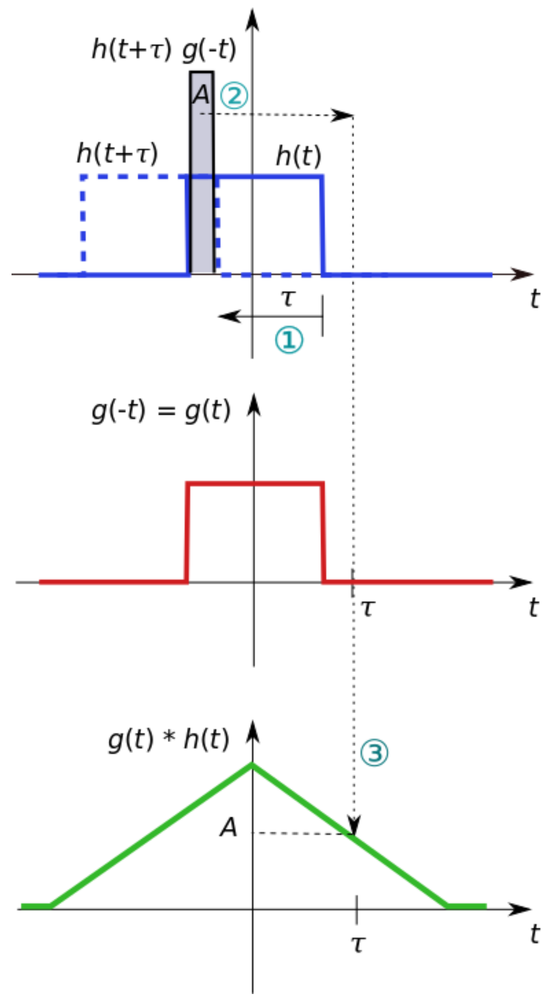

Figure 2: The convolution between two input functions $g(t)$ and $h(t)$ in the specific case of $g(t)$ being even.

First, we rotate function $g(t)$ around the vertical axes, which has no effect in this particular case as $g(t)$ is an even function. Then, for each value of $-\infty < \tau < \infty$, a shifted version of $h(t)$, namely $h(t+\tau)$ is obtained and multiplied by $g(-t)$, yielding a new function, shown in gray, whose area A corresponds to $[g * h](\tau)$ for each specific value of τ.

The more general situation in which $g(t)$ is not even can be immediately addressed by rotating this function around the vertical axis as the first step.

Observe that the convolution operation tends to increase the extension along t (non-zero values) of the resulting function with respect to the two original inputs. Actually, it can be shown that this extension corresponds to the sum of the extensions of the $g(t)$ and $h(t)$ in the case these extensions are bound.

This example illustrates why convolution is sometimes understood as a blend of its two input functions $g(t)$ and $h(t)$, in the sense that $[g * h](\tau)$ inherits characteristics from both original functions.

3 The Dirac Delta 'Function'

Formally speaking, the Dirac delta 'function' $\delta(t)$ is not a function as its valued is not defined at $t = 0$. According to the theory of distributions (e.g. [2]), this function corresponds to a *functional* that takes an original, 'well-behaved' function $g(t)$ into the scalar corresponding to its value at zero, i.e. $g(0)$.

Consider the rectangular function with area 1 given as

$$r(t) = \begin{cases} \mathbf{c} = \mathbf{\frac{1}{2a}}, & \text{for } t \in [-a, a] \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

The Dirac delta 'function' can be informally understood as the following limit of $r(t)$

$$\delta(t) = \lim_{a \to \infty} r(t) \tag{5}$$

Since $r(t)$ has unit area for any $a > 0$, it follows that

$$\int_{-\infty}^{\infty} \delta(t) dt = 1 \tag{6}$$

We will informally understand that

$$\delta(t)\, g(t) = \delta(t)\, g(0) \tag{7}$$

and we obtain

$$\int_{-\infty}^{\infty} \delta(t) g(t) dt = g(0) \tag{8}$$

which is sometimes understood as the *sampling* property of the Dirac delta 'function'.

This result allows us to easily obtain the convolution of a function $g(t)$ with the Dirac delta, i.e.

$$\int_{-\infty}^{\infty} \delta(-t) g(t+\tau) dt = \int_{-\infty}^{\infty} \delta(-t) g(\tau) dt = g(\tau) \tag{9}$$

Interestingly, this results in the original function $g(t)$, and we can understand the Dirac delta as a kind of *identity element* of the convolution. Figure 3 illustrates this result graphically.

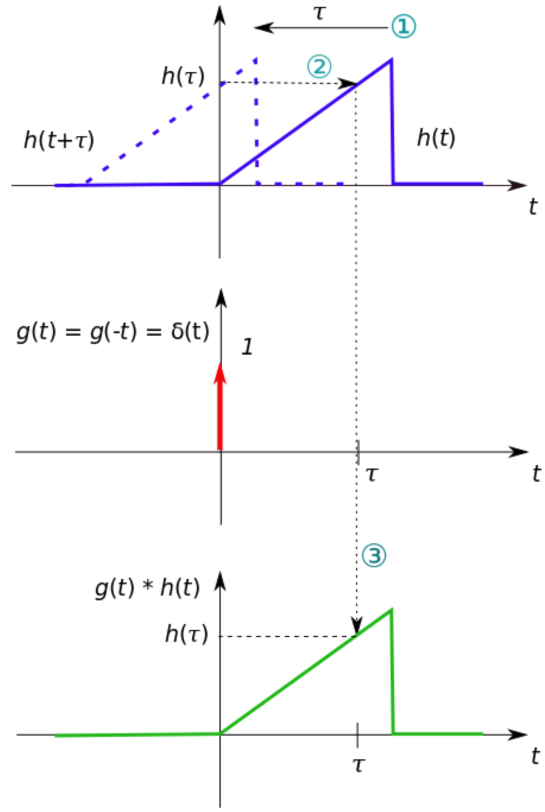

Figure 3: The convolution between a given function $h(t)$ and the Dirac delta results in the function $h(t)$ placed at the position of the Dirac delta.

A relatively frequent situation involves convolving a function $g(t)$ with a sum of Dirac deltas, such as

$$S(t) = \sum_{k=-\infty}^{\infty} \delta(t - k\,\Delta t) \tag{10}$$

where Δt is the spacing between the infinite sequence of Dirac deltas.

The effect of convolving $S(t)$ with a function $g(t)$ can be easily understood by taking into account that the convolution is a bilinear operation, from which follows that

$$[S * g](\tau) = \sum_{k=-\infty}^{\infty} g(t - k\,\Delta t) \tag{11}$$

In other words, this convolution involves *adding* $g(t)$ at each of the positions $k\Delta t$ at which the Dirac deltas are located.

4 Convolution as 'Blending' of Two Functions

Convolution can often be understood as performing some kind of 'blending' or 'mixing' between the two involved functions $g(t)$ and $h(t)$, in the sense that the resulting function 'inherits' features from both $g(t)$ and $h(t)$.

In order to understand this type of interpretation, let's start by observing that the convolution of any function $g(t)$ with the constant function $h(t) = c$ yields another constant function, namely $z(t) = c \int_{-\infty}^{\infty} g(t)dt$ (i.e. the area of $g(t)$), as a result.

Let's also start by considering the *sinc* function as being defined as:

$$sinc(t) = \begin{cases} \frac{sin(\pi t)}{\pi t}, & \text{for } t \neq 0 \\ 1, & \text{for } t = 0 \end{cases} \tag{12}$$

Observe that this is the normalized version (area equal to 1) of the *sinc* function, otherwise defined as $sin(t)/t$.

The sinc is an example of a function *localized* along the time domain, in the sense that its 'mass' is concentrated at some point R (in this case 0), as can be inferred considering the mean and standard deviation of the sinc function. Informally speaking, a localized function will have most of its area ('mass') concentrated around the interval having a few standard deviations as width and centered at R. Being localized also implies that the function tends to decrease to the left and righthand sides R.

Let's also consider the *rectangle function* with width L as being defined as:

$$R_L(t) = \begin{cases} 1, & \text{for } -\frac{L}{2} < t \leq \frac{L}{2} \\ 0, & \text{otherwise} \end{cases} \tag{13}$$

If L is much larger than the standard deviation of the sinc (or any other localized function), we have that, except the regions near the transitions of $R_L(t)$, the convolution between the sinc ($g(t)$) and the rectangle function ($h(t)$) can be well approximated by a convolution between the sinc and a constant function, so that the it results approximately to the sinc area $\int_{-\infty}^{\infty} g(t)dt$, which we know by definition to be equal to 1. So, convolving the sinc over the rectangle function will have almost no effect in the middle portion of the latter function, which will remain almost unaltered.

However, the convolution will produce different results around each of the left and right transitions of the rectangle functions. Figure 4 illustrates the convolution between a sinc (a) and a wider rectangle function (b). The resulting function (c) can be understood as a 'blend' or 'mix' of the two original functions, inheriting features from both of them.

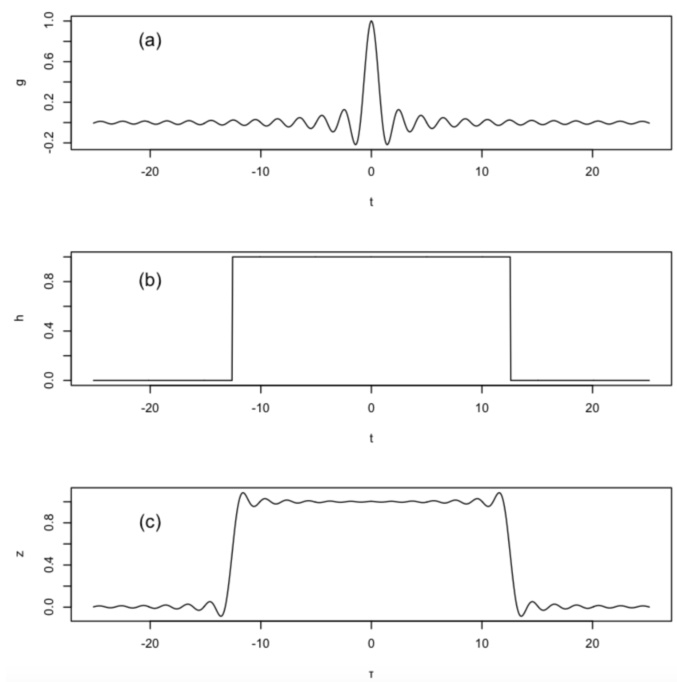

Figure 4: The convolution between a localized function (a) and a wider function (b) tends to yield a new function (c) that can be understood as a 'blend' or 'mixing' of the two original functions, in the sense of inheriting features from both of them.

5 Convolution as 'Matching' of Two Functions

When two localized functions have similar sizes, such as in Figures 5(a) and (b), the convolution between them tends to provide a function (c) that peaks at the time position corresponding to the best 'matching' or 'alignment' between the two original functions. Usually, a better 'matching' can be obtained by using the correlation (see Section 12) instead of convolution, since this binary operation does not involve mirroring of any of the two original functions.

The 'matching' feature of the convolution is related to the concept of *internal product* between two real functions $g(t)$ and $h(t)$, which is typically defined as:

$$\langle g(t), h(t) \rangle = \int_{-\infty}^{\infty} g(t)h(t)dt \tag{14}$$

Observe that, in the case of complex functions $g(t)$ and $h(t)$, the internal product typically involves the conjugation of $h(t)$.

The interpretation of this important binary operation can be more conceptually approached by considering its *discrete* version, i.e. considering finite versions of $g(t)$ and $h(t)$ sampled along time and within a time window, giving rise to two respective real vectors $\vec{g}$ and $\vec{h}$ with N elements each. In this case, the internal product is given as:

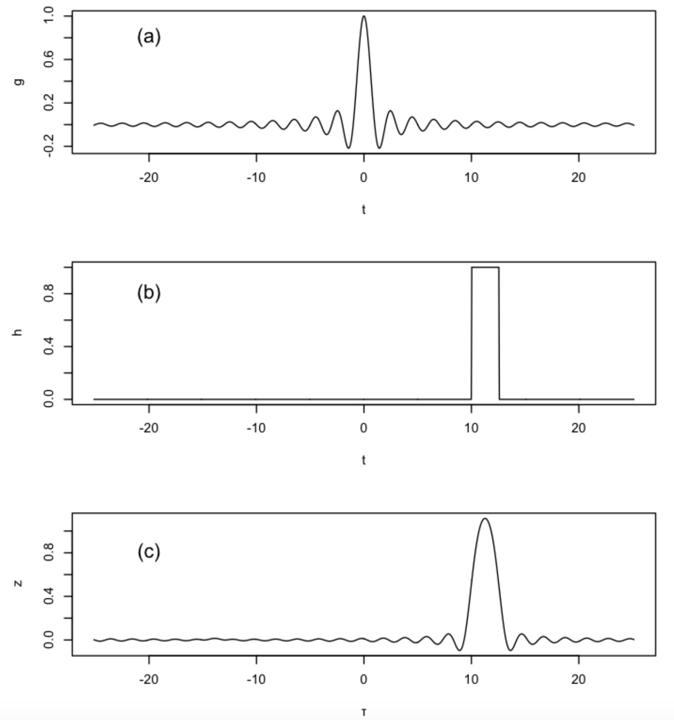

Figure 5: The convolution between two localized functions with similar widths (a,b) tends to yield a new function (c) peaking at the time position corresponding to the best 'matching' or 'alignment' between the two original functions.

$$\left\langle \vec{g}, \vec{h} \right\rangle = \sum_{i=1}^{N} g_i \, h_i = |\vec{g}||\vec{h}|cos(\theta) \tag{15}$$

where θ is the smallest angle between the two vectors. It follows from this definition that, while keeping $|\vec{g}|$ and $|\vec{h}|$ vixed, the values of the internal product will depend on θ, reaching its highest value for $\theta = 0$, minimal (negative) value for $\theta = \pi$, and null value for $\theta = \pi/4$. So, the internal product between $\vec{g}$ and $\vec{h}$ can be conceptually understood as a measurement of the similarity between the angles of these vectors (remember that we chose to keep their magnitudes constant). This interpretation can be somewhat extended to the internal product between the two original functions $g(t)$ and $h(t)$ from which the vectors $\vec{g}$ and $\vec{h}$ were derived.

Now, we can go back to the convolution definition(Eq. 3) and observe that it actually corresponds to the internal product between $g(-t)$ and $h(\tau + t)$. As such, we can conceptually interprete that, for functions with fixed magnitude, the convolution at a given value of relative displacement τ provides some measurement of the 'similarity between $g(-t)$ and $h(t)$. That is, ultimately, the reason why the convolution can be understood as a 'matching' between the two involved real functions. It also explains why, usually, a better 'matching' can be provided by the correlation, which does not invert one of the functions along time, therefore corresponding more directly to the internal product between $g(t)$ and $h(t)$. In case these functions are even, there will be no difference in inferring the 'matching' between them by using either correlation or convolution.

Observe that special care needs to be taken regarding the magnitudes of the involved functions, which were assumed constant in the above discussion. In case the magnitudes can vary, we need to take their direct effect on the 'matching' intensity.

6 The Fourier Transform

The Fourier transform is intrinsically related to convolution and the Dirac delta 'function', and as such it allows further insights to be obtained regarding the meaning of convolution. In this section we briefly present the continuous Fourier transform in one dimension (e.g. [3]).

Let $g(t)$ be a complex (or real) function of t. Under certain circumstances (e.g. [3]), its Fourier transform exists and can be calculated as

$$\mathcal{F}\{g(t)\} = G(f) = \int_{-\infty}^{\infty} g(t) exp\{-i2\pi f t\} \, dt \tag{16}$$

It is also possible to adopt $\omega = 2\pi f$.

Observe that this transform maps the original function $g(t)$ in the 'time' domain into a new complex function $G(f)$ in the 'frequency' domain. The original function $g(t)$ can be recovered (under some conditions) from $G(f)$ as

$$\mathcal{F}^{-1}\{G(f)\} = g(t) = \int_{-\infty}^{\infty} G(f) exp\{i2\pi f t\} \, df \tag{17}$$

When both the above direct and inverse Fourier transforms exist, we have a respective *Fourier transform pair*.

$$g(t) \longleftrightarrow G(f) \tag{18}$$

A first interesting property of the Fourier transform, which follows easily from its definition, is that this transform is *linear*, i.e. $\mathcal{F}\{a\, g(t) + b\, h(t)\} = a\, \mathcal{F}\{g(t)\} + b\, \mathcal{F}\{h(t)\}$.

Let's calculate the Fourier transform of $\delta(t)$.

$$\mathcal{F}\{\delta(t)\} = \int_{-\infty}^{\infty} \delta(t) exp\{-i2\pi f t\} \, dt =$$
$$= \int_{-\infty}^{\infty} \delta(t) exp\{-i2\pi f(0)\} \, dt = \int_{-\infty}^{\infty} \delta(t) dt = 1 \tag{19}$$

It can be shown that the inverse Fourier transform of 1 exists and corresponds to $\delta(t)$, so we can write our first *Fourier transform pair* as

$$\delta(t) \leftrightarrow 1 \tag{20}$$

It can be verified that the Fourier transform of a purely real and even function $g(t)$ is real and even. Similarly, the Fourier transform of a purely imaginary and odd function is imaginary and odd.

The Fourier transform has a particularly interesting, though not so often applied, property known as *symmetry* (e.g. [3]). If we have the Fourier pair $g(t) \leftrightarrow G(f)$, it follows that

$$G(t) \longleftrightarrow g(-f) \tag{21}$$

This property allows us to immediately derive several new Fourier transform pairs from already known results. For instance, if we apply this property to the pair $\delta(t) \leftrightarrow 1$, we conclude that

$$1 \leftrightarrow \delta(f) \tag{22}$$

Let's now see what happens when we shift a function $g(t)$ to the left by t_0, i.e. $g(t-t_0)$. From the Fourier transform definition, we have that

$$\mathcal{F}\{g(t-t_0)\} = \int_{-\infty}^{\infty} g(t-t_0) exp\{-i2\pi f t\}\, dt \tag{23}$$

now, we make the variable transformation $u = t - t_0$, which implies

$$\mathcal{F}\{g(t-t_0)\} = \int_{-\infty}^{\infty} g(u) e^{-i2\pi f(u+t_0)} du =$$
$$= e^{-i2\pi f t_0} \int_{-\infty}^{\infty} g(u) e^{-i2\pi f u} du = e^{-i2\pi f t_0} G(f) \tag{24}$$

It can be shown that this *time shifting* property also holds with respect to the inverse Fourier transform, so that we obtain the pair

$$g(t-t_0) \longleftrightarrow e^{-i2\pi f t_0} G(f) \tag{25}$$

The *frequency shifting* property can be similarly derived as

$$e^{i2\pi t t_0} g(t) \longleftrightarrow G(f - f_0) \tag{26}$$

At this point, we can obtain several interesting Fourier pairs involving sinusoidals and complex exponentials. We start with the derivation of the Fourier pair for $e^{i2\pi f t}$. By combining the fact that $1 \leftrightarrow \delta(f)$ with the frequency shift property, it follows that

$$e^{i2\pi f_0 t} \longleftrightarrow \delta(f - f_0) \tag{27}$$

Given that $cos(2\pi f_0 t) = 1/2\left(e^{i2\pi f_0 t} + e^{-i2\pi f_0 t}\right)$, we immediately find that

$$cos(2\pi f_0 t) \longleftrightarrow \frac{1}{2}\left[\delta(f+f_0) + \delta(f-f_0)\right] \tag{28}$$

Analogously, it follows from $sin(2\pi f_0 t) = 1/(2i)\left(e^{i2\pi f_0 t} - e^{-i2\pi f_0 t}\right)$ that

$$sin(2\pi f_0 t) \longleftrightarrow \frac{1}{2i}\left[-\delta(f+f_0) + \delta(f-f_0)\right] \tag{29}$$

hence

$$sin(2\pi f_0 t) \longleftrightarrow \frac{i}{2}\left[\delta(f+f_0) - \delta(f-f_0)\right] \tag{30}$$

We also have, from the time shifting property, that

$$\delta(t-t_0) \longleftrightarrow e^{-2\pi f t_0} \tag{31}$$

So, if we imagine that the original function $g(t)$ is composed by a (possibly infinite) sum of properly time shifted Dirac deltas $\delta(t-t_0)$, we obtain in the frequency domain a respective sum of above complex exponentials (involving sines and cosines). When f is discretized, function $g(t)$ becomes periodic and the inverse Fourier transform becomes related to the Fourier *series*.

This fact explains why the sine (and cosine) functions are particularly important for the Fourier transform, defining its respective *basis*, which is used to express the function $g(t)$. The recover of $g(t)$ is performed by the inverse Fourier transform in terms of typically infinite linear combinations of sines and cosines, which we will call *harmonic components*.

Table 1 lists some commonly used pairs of Fourier transforms.

7 The Convolution Theorem

One of the strong connections between convolution and the Fourier transform is established through the convolution theorem, which states that

$$g(t) * h(t) \longleftrightarrow G(f)H(f) \tag{32}$$

We also have that

$$g(t)\, h(t) \longleftrightarrow G(f) * H(f) \tag{33}$$

So, the Fourier transform allows convolutions in the time domain to be performed in terms of function pointwise multiplications in the frequency domain, and vice versa. This can often allow not only substantial savings of computational expenses, but also contribute to a more complete conceptual understanding of the convolution, as developed in the following section.

Table 2 lists some other commonly used Fourier transform properties.

8 Convolution in the Frequency Domain

We are now in a position to consider convolution between any pair of sinusoidal functions, such as two cosines. It is relatively difficult to infer the result of the convolution of these two functions by considering their distribution along time, e.g. by using graphical constructions such as that described in Section 2.

Figure 6 shows two such functions, with respective frequencies $f_1 = 1kHz$ and $f_2 = 2kHz$, as well as their respective Fourier transforms (recall that, as a consequence of evenness of the cosine functions, we have purely real Fourier transform results).

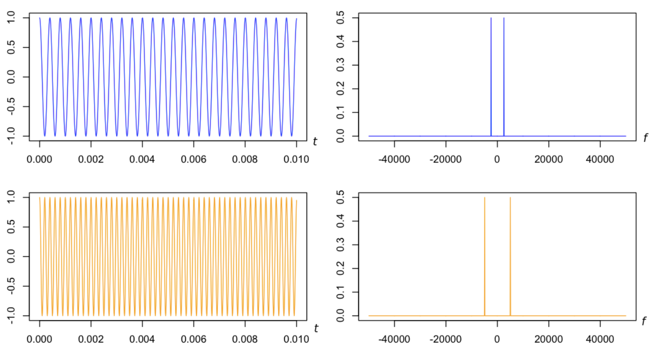

Figure 6: A function $g(t)$ corresponding to the sum of two cosine functions with distinct respective frequencies $f_1 = 1kHz$ and $f_2 = 2kHz$ and its Fourier transform.

We have from the convolution theorem that this convolution between two cosines can be obtained as the inverse Fourier transform of the product of their respective transforms. As a consequence of $f_1 \neq f_2$, we have that this product in the frequency domain results in the null function (0 everywhere), so that the respective inverse Fourier is also null. We conclude that the convolution between two cosines (or sines) will only be non-zero in case they both have the same frequency.

This interesting result immediately extends any two functions defined by finite sums of sines and/or cosines (i.e. periodic functions, related to the Fourier series).

Figure 7 illustrates a situation with respect to two functions containing only one ($f_1 = 1kHz$) and two ($f_1 = 1kHz$ and $f_2 = 2kHz$) harmonic components, respectively.

Now, the components respective to $f_1 = 1kHz$ in both functions align in the frequency domain, resulting in a non-null convolution outcome that contains only this frequency, though with changed amplitude.

So, the consideration of the convolution action in the frequency domain, as allowed by the Fourier transform, can contribute to our better understanding of the nature and effects of the convolution operation of periodic func-

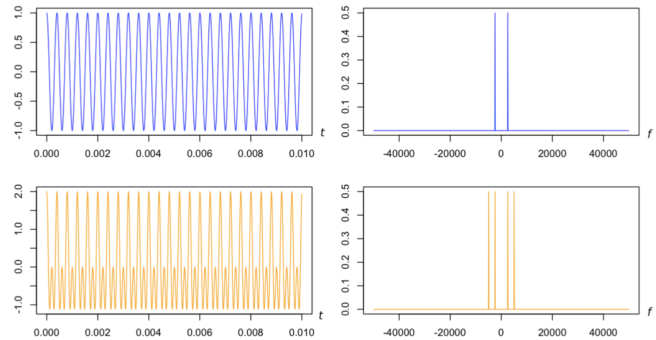

Figure 7: Two functions defined by sums of cosines, sharing one of the respective frequencies, and the respective Fourier transforms.

tions, including sines and cosines.

9 Linear, Time Invariant Systems

Let's now consider an interesting application of the convolution as a powerful means to accurately model linear, time invariant dynamical systems. Figure 8 illustrates the typical graphic representation of a dynamic system, involving respective input $x(t)$ and output $y(t)$. The respective Fourier transforms $X(f)$ and $Y(f)$ are also shown.

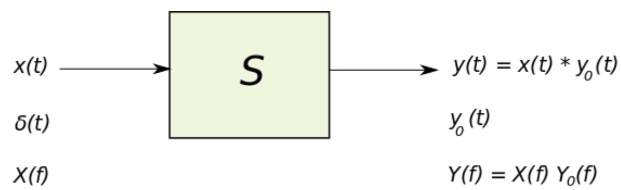

Figure 8: A linear, time invariant dynamic system, including its input and output represented in the time and frequency domain. The impulse response corresponds to $y_0(t)$, while $Y_0(f)$ is often called frequency response or transfer function.

This system will be *linear* iff $S(a\, x_1(t) + b\, x_2(t)) = a\, S(x_1(t)) + b\, S(x_2(t))$.

We also consider that system S is *time-invariant*, i.e. if $y(t) = S(x(t))$ then $y(t - t_0) = S(x(t - t_0))$ for any $t_0 \in \mathbb{R}$. More informally, this means that the properties of S do not change along time.

Let's also define the *impulse response* of the system S as the function $y_0(t)$ obtained when we input a Dirac delta $x(t) = \delta(t)$ (the *impulse*) into S (also shown in Figure 8).

Under these circumstances, it can be shown that, for any input function $x(t)$, we have that

$$y(t) = x(t) * y_0(t) \tag{34}$$

So, once the system has been probed with the impulse $\delta(t)$, yielding the respective impulse response $y_o(t)$, its respective operation considering any other input can be provided by the above convolution. In other words, all dynamical properties of S are captured into $y_0(t)$, which therefore assimilates every possible information about the inner workings of S.

This interesting — and useful — property, which may initially sound like 'magic, can be easily understood in an intuitive way in the frequency domain. First we express the expression in Equation 34 in the frequency domain

$$Y(f) = X(f)\, Y_0(f) \tag{35}$$

Now, remember that the Fourier transform of the Dirac delta impulse is the constant function $h(t) = 1$. In other words, this function will equally stimulate every possible frequency inside the system S, as it modifies each of these frequencies in its specific way, yielding the resulting *frequency response*, or *transfer function* $Y_0(f)$.

So, we can understand the system operation as being defined by the $Y_0(f)$ and, as a consequence, the response to any other input function $x(t)$ can now be obtained as the inverse Fourier transform of the signal $X(f)\, Y_0(f)$.

10 Linear Filters

The above developed framework to model linear, time-invariant systems can be used to introduce the concept of *linear filters*. Basically, we consider the filtering function as corresponding to the transfer function above, which is kept fixed during the system operation which will, therefore, act on the signal $X(f)$ modifying it as specified by the filter function $Y_o(f)$,.

Linear filters are frequently used in a wide range of applications, such as removing additive noise, change the tone of sound signals (bass/treble), communications, emphasizing signal transitions, etc.

Basically, the operation of the filter $Y_0(f)$ is specified by the distribution of its values along f. Figure 9 illustrates four basic types of filters, namely: (a) *low-pass*, (b) *high-pass*, (c) *band-pass* and (d) *band-rejection* (or *band-stop*).

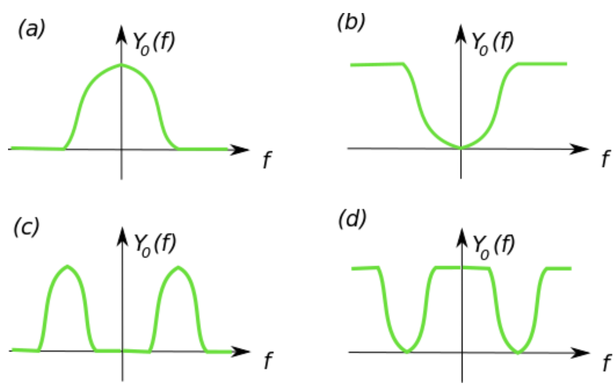

Figure 9: Four types of commonly considered filters: (a) *low-pass*, (b) *high-pass*, (c) *band-pass* and (d) *band-rejection* (or *band-stop*).

It is important to keep in mind that the terms 'high' and 'low' typically refer to the absolute value of the involved frequencies. Observe that Fourier transforms normally exted along both sides of the frequency axis.

Therefore, *low-pass filters*, allowing the lower frequencies to pass while attenuating or eliminating the higher frequencies, involve a filter function having higher (positive or negative) values near the frequency axis origin. High-pass filters have an opposite nature. Band and rejection-pass filters allow only an interval of frequencies to pass, reducing or blocking the other frequencies.

11 Deconvolution

We have seen how a function $x(t)$ can be modified by convolving it with another function $y(t)$ (e.g. a filter), resulting a function $z(t)$ or, in the frequency domain $Z(f) = X(f)Y(f)$.

It is often necessary to consider how $x(t)$ can be recovered from $z(t)$. Provided $Z(f)$ do not assume any zero value at any frequency f, we can use the convolution theorem to obtain

$$X(f) = \frac{Z(f)}{Y(f)} \tag{36}$$

This deterministic procedure is called *deconvolution*, as the means of reversing the convolution effects.

In practice, we often do not known $Y(f)$, which therefore has to be estimated somehow. In the case of experimental signals, we often will also have some involved noise signal $\epsilon(t)$, such as in the following case of *additive* noise

$$z(t) = x(t) * y(t) + \epsilon(t) \tag{37}$$

The procedure of determining $y(t)$ in these circumstances is the subject of interest of the area of *statistical signal processing*, involving methods such as the Wiener and Kalman filters (e.g. [4]).

12 Correlation

Correlation corresponds to another interesting operation similar to the convolution that also finds several applications, such as in statistics (self- and cross-correlations) as well as in pattern matching or recognition. The (cross-)correlation between two complex functions $g(t)$ and $h(t)$ is typically defined as

$$g(t) \otimes h(t) = \int_{-\infty}^{\infty} g(t)^* h(t-\tau) dt \tag{38}$$

where $*$ means complex conjugation.

The alternative convolution integral adopted in the present work allows a direct comparison between convolution and correlation. Comparing Equations 3 and 38, we observe that convolution and correlation are very similar except for the negation of t in $g(t)$ and the the fact that $h(t)$ is shifted in opposite direction.

Correlation has a property analogous to the convolution theorem:

$$g(t) \otimes h(t) \longleftrightarrow G^*(f)\, H(f) \tag{39}$$

This is often called the *correlation theorem*.

Given a signal $g(t)$, its *autocorrelation* can be obtained in the frequency domain as

$$G^*(f)\, G(f) = |G(f)|^2 \tag{40}$$

which corresponds to the *power spectrum* of $g(t)$.

The *energy* in $g(t)$ can be understood as corresponding to

$$E_g = \int_{-\infty}^{\infty} |g(t)|^2 dt \tag{41}$$

Parseval's theorem indicates that the energy of the original signal $g(t)$ is preserved in the respective frequency representation $G(f)$, i.e.

$$E_g = \int_{-\infty}^{\infty} |g(t)|^2 dt = \int_{-\infty}^{\infty} |G(f)|^2 df = E_G \tag{42}$$

13 Concluding Remarks

Convolution plays a particularly general and important role in many scientific and technological areas, from physics to dynamic systems. Yet, given its relatively intricate complex integral definition, it is sometimes not so well understood intuitive or conceptually. In this present text, we attempted at providing a smooth introduction to convolution, possibly leading to a better and more comprehensive understanding of this important concept.

In order to do so, we resorted to an alternative integral definition of the convolution where the time variable appears without negation, allowing a potentially simpler and more intuitive understanding and graphical representation. In addition, the convolution was discussed as a means to 'blend' or 'mix' the two original function (usually when these are localized along time and one of them is wider than the other), and also as a means to identify the time position where the two original function best 'match' or 'alignt' one another (typically when the two functions are localized and have similar widths). We also considered the close relationship between convolution and the Fourier transform and the convolution theorem, which allowed us to understand that convolution between functions containing a finite number of discrete harmonic components often yield the null function as result.

The developed concepts also allowed a straightforward presentation of the important concepts as the impulse response and filters in linear, time-invariant dynamic systems. The concepts of deconvolution and correlation were also briefly introduced.

In a universe permeated by 'waves' of complexity (e.g. [5]), convolution represents a relatively simple concept paving the way to effectively mixing waves in the most diverse range of possible applications.

Acknowledgments.

Luciano da F. Costa thanks CNPq (grant no. 307085/2018-0) for sponsorship. This work has benefited from FAPESP grant 15/22308-2.

Costa's Didactic Texts – CDTs

CDTs intend to be a halfway point between a formal scientific article and a dissemination text in the sense that they: (i) explain and illustrate concepts in a more informal, graphical and accessible way than the typical scientific article; and (ii) provide more in-depth mathematical developments than a more traditional dissemination work.

It is hoped that CDTs can also integrate new insights and analogies concerning the reported concepts and methods. We hope these characteristics will contribute to making CDTs interesting both to beginners as well as to more senior researchers.

Though CDTs are intended primarily for those who have some experience in the covered concepts, they can also be useful as summary of main topics and concepts to be learnt by other readers interested in the respective CDT theme. Observe that CDTs come with absolutely no warranty.

Each CDT focuses on a few interrelated concepts. Though attempting to be relatively self-contained, CDTs also aim at being relatively short. Links to related material are provided in order to complement the covered subjects.

The complete set of CDTs can be found at: `https://www.researchgate.net/project/Costas-Didactic-Texts-CDTs`.

Table 1: Some Fourier transform pairs $g(t) \leftrightarrow G(f)$. $H(t)$ is the heavyside step function, $sng(t) = 2H(t) - 1$ is the sign function, R_L is the rectangle function with width L, and $\Delta t = 1/f_0 \in$ R.

$g(t)$	$G(f)$
1	$\delta(f)$
$\delta(t)$	1
$e^{i2\pi f_0 t}$	$\delta(f - f_0)$
$e^{-a\|t\|}$	$\frac{2a}{a^2+(2\pi f)^2}$
$cos(2\pi f_0 t)$	$\frac{1}{2}\left[\delta(f+f_0) + \delta(f-f_0)\right]$
$sin(2\pi f_0 t)$	$\frac{i}{2}\left[\delta(f+f_0) - \delta(f-f_0)\right]$
$H(t)$	$\frac{1}{2}\delta(f) + \frac{1}{i2\pi f}$
$R_L(t)$	L sinc(L t)
$sgn(t)$	$\frac{1}{i2\pi f}$
e^{-at^2}	$\frac{\sqrt{\pi}}{a} e^{-\pi^2 f^2/a}$
$\sum_{n=-\infty}^{\infty} \delta(t - n\Delta t)$	$\frac{1}{\Delta t} \sum_{n=-\infty}^{\infty} \delta(t - \frac{n}{\Delta t})$

Table 2: Some properties of the Fourier transform, assuming $g(t) \leftrightarrow G(f)$, $h(t) \leftrightarrow H(f)$, and $a, b \in \mathbb{C}$.

$a\, g(t) + b\, h(t) \longleftrightarrow a\, G(f) + b\, H(f)$
$G(t) \longleftrightarrow g(-f)$
$g(-t) \longleftrightarrow G(-f)$
$e^{i2\pi f_0 t} \longleftrightarrow G(f - f_0)$
$g(t - t_0) \longleftrightarrow e^{-i2\pi t_0 f} G(f)$
$g(a\, t) \longleftrightarrow \frac{1}{\|a\|} G(\frac{f}{a})$
$g(t) * h(t) \longleftrightarrow G(f) H(f)$
$g(t) h(t) \longleftrightarrow G(f) * H(f)$
$g(t) \otimes h(t) \longleftrightarrow G^*(f) H(f)$
$\frac{d^a}{dt^a} g(t) \longleftrightarrow (j2\pi f)^a G(f)$
$\int_{-\infty}^{t} g(\tau) d\tau \longleftrightarrow \frac{G(f)}{(j2\pi f)} + \frac{1}{2} G(0) \delta(f)$
$\int_{-\infty}^{\infty} \|g(t)\|^2 dt = \int_{-\infty}^{\infty} \|G(f)\|^2 df$

References

[1] L. da F. Costa. Group theory: A primer. Researchgate, 2019. `https://www.researchgate.net/publication/334126746_Group_Theory_A_Primer_CDT-11`. Online; accessed 30-July-2019.

[2] G. van Dijk. *Distribution Theory*. De Gruyter Graduate Lectures, 2013.

[3] E. O. Brigham. *Fast Fourier Transform and its Applications*. Pearson, 1988.

[4] R. G. Brown and P. Y. C. Hwang. *Introduction to Random Signals and Applied Kalman*. Wiley and Sons, 1996.

[5] L. da F. Costa. Quantifying complexity. Researchgate, 2019. `https://www.researchgate.net/publication/332877069_Quantifying_Complexity_CDT-6`. Online; accessed 30-July-2019.