The Unit Impulse Function

a natural understanding

K Poornima Guruprasad Ullal Rajeev Rajkumar Stephen Vadakkan

Department of Computer Science & Engineering

NMAM Institute of Technology

Nitte, Udupi District

Karnataka- 574110

INDIA

Disclaimer

This explanation is to give the beginner a feel for the Unit Impulse Function: the connection between what happens in nature and the related mathematics. It does not purport to be correct. In particular pay close attention to the mathematical expressions/formulae on page 14.

In the standard definition of the Unit Impulse Function there is the Mathematical oddity

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

This is equivalent to saying: $\infty \bullet$ ZERO $= 1$. This explanation avoids this oddity. Hope the reader finds it useful. All comments and criticism are welcome.

We do not accept any responsibility for any loss by the reader due to interpretation or application or any other reasons.

Let us try to understand the Unit Impulse Function by looking at a real life natural example.

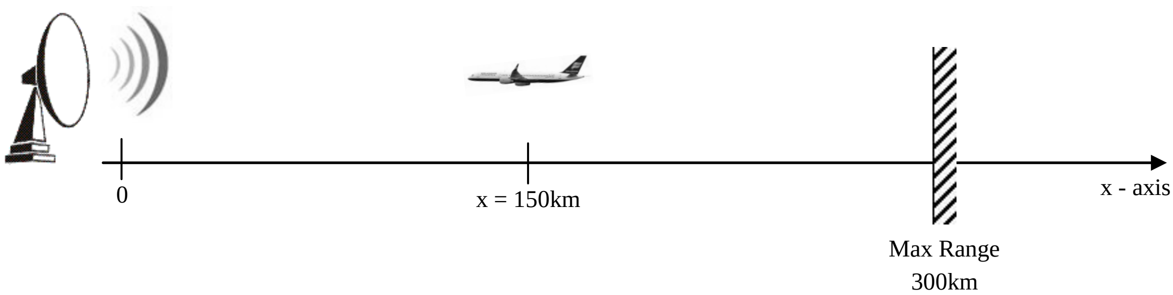
We may determine the exact distance to a target by flashing a beam of light on the target at instant t_1 and noting the instant t_2 at which the reflected beam returns to the source.

Then distance to target = (c = speed of light) • $(t_2 - t_1) / 2$

This is the basic principle behind radar.

Instants and exact distances are theoretical concepts. Also the range of a beam has its limitations.

Lets us now work with a radar that has a maximum range of 300km.



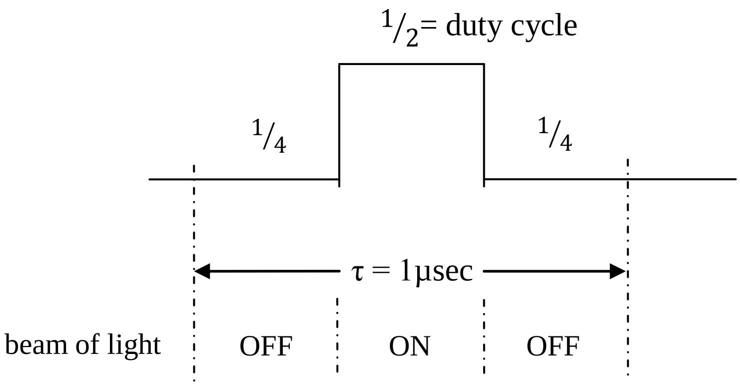
The time taken to max range and back is

$$\frac{\text{Twice Max. Range}}{\text{C = Speed of Light}} = \frac{2(3 \text{ x } 10^5 \, meters)}{3 \text{ x } 10^8 \, m/sec} = \frac{2}{1000}\text{sec}$$

$$= 2\text{millisec}$$

$$= 2000\mu\text{sec}$$

Since in practice we cannot work with instants, let us choose a clock pulse $\tau = 1$ µsec = δt.



$T_{duty} = T_d = {}^1/_2\mu\text{sec}$

$T_{pulse} = T_p = 1\mu\text{sec}$

We set the clock pulse counter to zero. We send/pulse a beam of light (for a duration of 1 clock pulse) as depicted above. We wait for up to 2000 μsec to sense an echo/response before sending/pulsing another beam of light.

Suppose, for example, we get a response after 1000 clock pulses. Then we can say that there is a target at $^{1000}/_{2}$ =500 clock pulses away.

So distance to target = 500 pulses X $\frac{1\mu\text{sec}}{pulse}$ X $\left(\frac{3 \text{ x } 10^{8}\ meters}{10^{6}\ \mu\text{sec}}\right)$

= 500 X 300 = 150,000 meters = 150kms

Because the duty cycle is $^{\tau}/_{2}$, the accuracy/range resolution is

Δ R = c . $^{\tau}/_{2}$ = 150 meters

Let us see another way to acquire this information. Let us look at sequences below:

Primes & their Prime Multiples	Sequences of Primes and Prime Powers
2	4, 6, 8, 10, 12, • • • 30, • • •, 60, • • •, 210, • • •
3	6, 9, 12, • • • 30, • • •, 60, • • •, 210, • • •
2^2 =4	8, • • • 12, • • • 60, • • •
5	10, 15, 30, • • •, 60, • • •, 210, • • •
7	

Given the sequences formed from 2, 3 and 5 we can find the LCM = 30. Vice versa, given the LCM = 30 we can find the sequences involved: the sequences are formed from the primes 2, 3 and 5.

We may apply this concept to a radar using several independent frequencies.

Let τ = 1μsec as before. So the highest frequency available to us is $f = \frac{1}{\tau} = 10^{6}$Hz = 1MHz. Since, c = f• λ, its wavelength λ = 300meters.

We shall use several frequencies (f =10^{6}Hz and less) to transmit and locate a target.

Let us assume that these frequencies are continuously beamed across the range. If a target appears, let us further assume that when an echo/response is sensed (at the point of transmission) we assume this happens when the peak of a wave strikes the target. This is what the picture looks like.

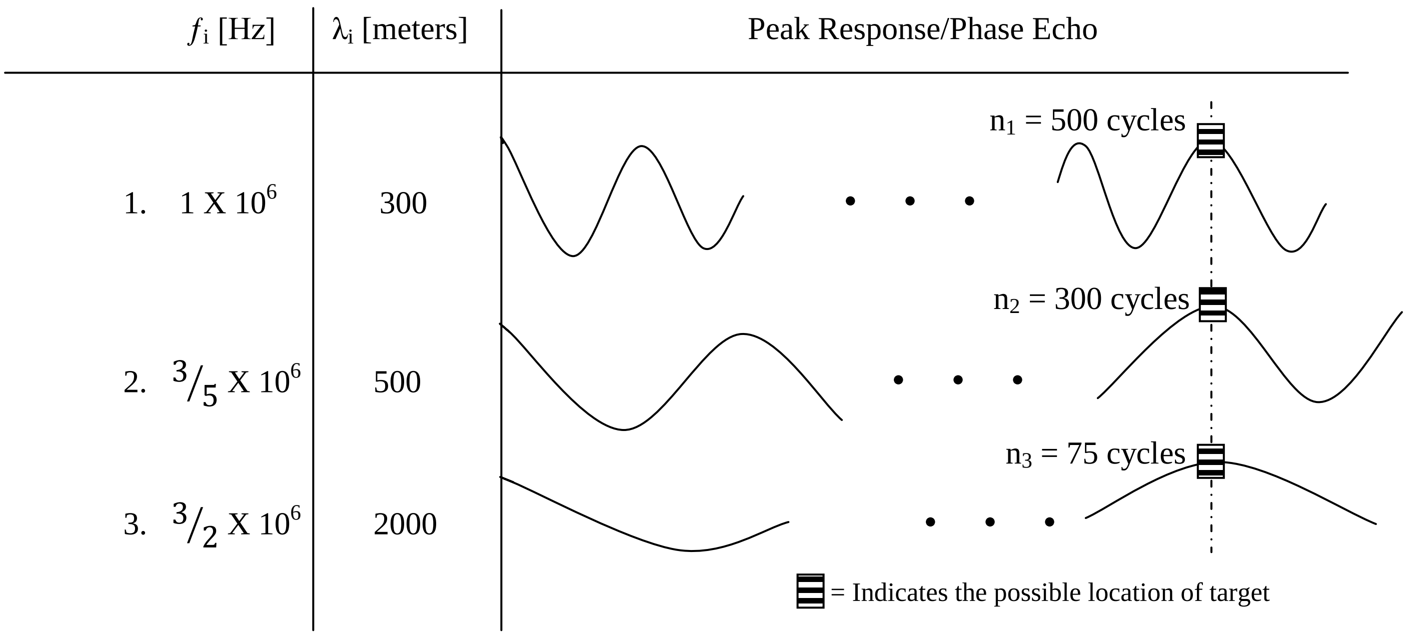


***Figure*: Spatial Domain**

On any single frequencies f_i , if there is an echo/response, the target may be on any of the many peaks in the range. Notice how each wave starts with a peak at t = 0. This is another way of saying the phase Ø = 0. Later we shall allow a peak to be at any position within a cycle, implying a non-zero phase or phase shift.

Moreover, when there are echoes from several frequencies we may infer that the target must be on peaks of these frequencies that line up/coincide as shown in the diagram.

Hence we have: $n_1\lambda_1 = n_2\lambda_2 = n_3\lambda_3 = \bullet\bullet\bullet$.

The calculation involved in locating the target with this information is analogous to finding the LCM given the sequences involved.

To match the example described, we can say with just two frequencies f_1 and f_2 that the target must be at the LCM of λ_1 = 300m and λ_2 = 500m and multiples of the LCM. The LCM is 1500. So the target may be at 1500m, 3000m, 4500, 6000m, • • • ,150000m, • • • .

How can we be MORE PRECISE about the location of the target?

Let us use our analogy of sequences. Suppose our target is actually located at 210 meters and we have only prime numbers 2, 3 and (highest frequency) 7 available to us as frequencies. Then the possible target locations are LCM (2, 3, 7) and its multiples: 42, 84, 126, 168, 210, 252, • • • .

We cannot uniquely determine the location. But with one additional prime number 5 we can uniquely determine the target location as LCM (2, 3, 5, 7) = 210.

So analogous to the above argument we can see that the more the frequencies present in our bandwidth from 0Hz to highest frequency10^6Hz the more precisely (i.e. fewer possible locations) we can determine the location of the target. With just three frequencies f_1, $f_2$and f_3 the target may be at the LCM (300m, 500m, 2000m) and its multiples. The LCM is 6000. So the target may be at 6000m, 12000m, 18000, 24000m, • • • ,150000m, • • • . We can see that with three frequencies the possible locations are fewer than with just two frequencies.

How can we guarantee that we can UNIQUELY determine the location of the target?

When we have all the frequencies from 0Hz to highest frequency10^6Hz we can uniquely determine the position of the target.

Note that using only FREQUENCY related information (without using any TIME related information such at τ = 1μsec and a count of clock pulses) we were able to locate the target.

What about the range resolution ΔR?

We know the highest/maximum frequency involved in 1MHz. It has wavelength 300 meters, which is the shortest of the frequencies involved. Because of the way we assume we get an echo/response, we may be off by half this wavelength, which is 150 meters. Thus the range accuracy is the same as when working with TIME related information. There will always be this error (range resolution) equal to half the (shortest) wavelength of the maximum frequency in the bandwidth.

How can we reduce this error and EXACTLY locate the target?

We can see that if we narrow the clock pulse we can reduce this error.

TIME INFORMATION	FREQUENCY INFORMATION
τ = 1μsec = 10^{-6} sec	max frequency f = 10^6 Hz = 1 MHz Bandwidth = 0→10^6 Hz λ = 300 meters
error = ½ τ.c = 150 meters	error = ½.λ = 150 meters
τ = 1nanosec = 10^{-9} sec	max frequency f = 10^9 Hz = 1 GHz Bandwidth = 0→10^9 Hz λ = 30cms
error = ½ τ.c = 15cms	error = ½.λ = 15cms
Ultimately: τ = instant	max frequency = ∞ Bandwidth = 0→ +∞ Hz λ = 0
error = 0	error = 0

So for locating the target on the +x-axis exactly (with zero error) we need a continuous set of frequencies from 0 to +∞. In the TIME DOMAIN the clock pulse becomes an instant. Another way to view this is: $T_d = T_p$ where $T_p > 0$.

The frequencies that return an echo are a subset of this continuous/complete set of frequencies.

How big is this subset? And, how are the frequencies in this subset related to each other?

With the understanding of discrete, dense and complete (continuous) in mind, we can see that this subset must be infinite. And with just high school knowledge of mathematics and physics (http://www.nitte.ac.in/downloads/Calculus.pdf) we can say these frequencies must be rational multiples of each other and they form a DENSE set. **In fact, this DENSE set is so DENSE that for practical purposes we may view it as a CONTINUOUS/COMPLETE set of frequencies.** Something like the Universal Guitar we saw in CONVOLUTION & LTI

Systems (http://www.nitte.ac.in/nmamit/articles.php?linkId=131&parentId=20&mainId=20&facId=131#).

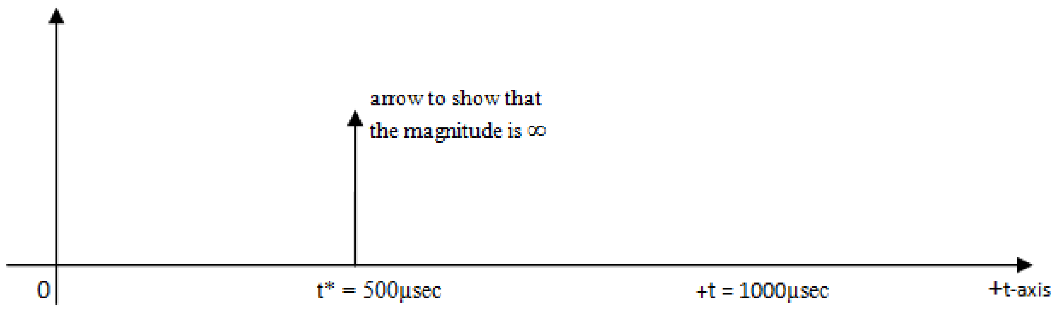
Let us now represent this SPATIAL DOMAIN information on the time axis or more precisely in the TIME DOMAIN.

The crucial information is the distance of the target = 150,000meters

This can be represented at exact instant $t^* = \frac{150{,}000 \text{ meters}}{3 \times 10^{8}(\text{meters} / 10^{6}\mu\text{sec})} = 500\mu\text{sec}$

What about the amplitudes of the frequencies (for which there is an echo)?

Each of these frequencies have amplitude = 1. The amplitude = 1 of each such frequency can be represented as a magnitude = 1 at t* = 500µsec. And since these frequencies combined identified the instant t* = 500µsec, we may add up these magnitudes to get ∞. Thus the TIME DOMAIN picture will look like



This is what we may call a UNIT IMPULSE at t* = 500µsec.

Let us now represent the SPATIAL DOMAIN information on the frequency axis or more precisely in the FREQUENCY DOMAIN.

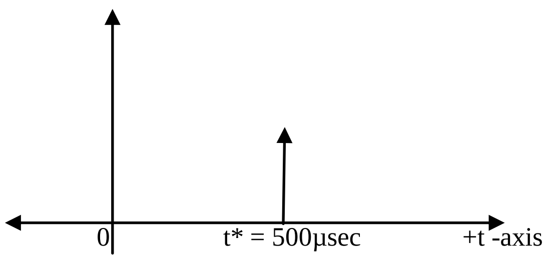
The picture will look like a continuous line because a dense set of frequencies are involved.

If the target shifts, say now by 100kms to the left from its current position of 150kms, a different DENSE set of frequencies will be involved in locating the target. Again we can represent this information with exact instant t* = 500µsec / 3.

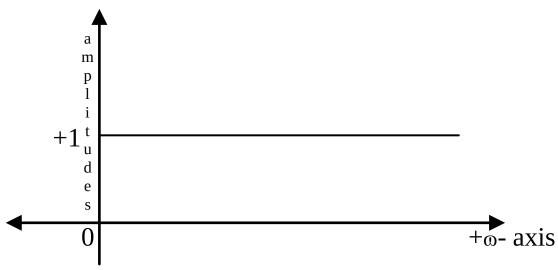
And again the magnitudes will add up to +∞. However the picture in the FREQUENCY DOMAIN looks the same because it is a DENSE SET (possibly a different DENSE SET) of frequencies. If we want the same DENSE SET then we should allow the phase to be non zero. This is analogous to saying 211 = 210 +1 where the phase shift of each prime power involved is 1. So we capture any instant in the TIME DOMAIN by the same DENSE SET of frequencies in the FREQUENCY DOMAIN. Later we shall see how we can capture a single frequency (whatever be the phase) in the FREQUENCY DOMAIN by a COMPLETE SET of instants in the TIME DOMAIN.

This is the picture no matter where the target is. So now we may relate the two pictures.

TIME DOMIAN

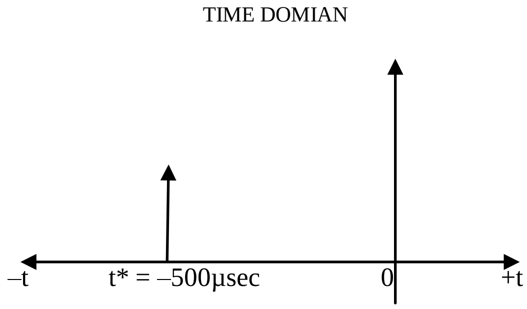


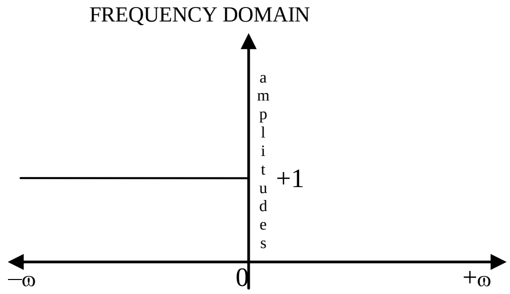
FREQUENCY DOMAIN



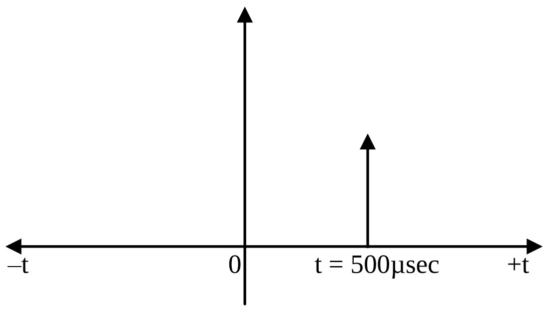
Let us now turn our radar in the negative x-axis direction.

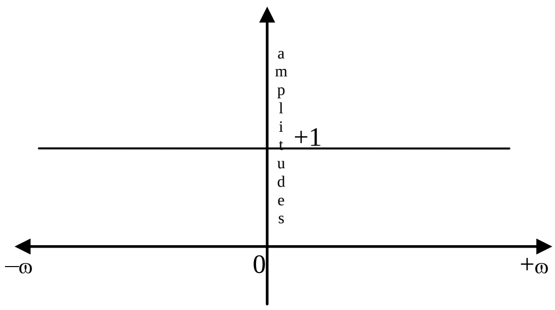
The same thinking and analysis applies. But this time with the waves propagating from 0 to –∞, that is to say in the negative direction. These waves have negative frequencies. So for a target located at –150kms the picture is





Now, with the target at +150kms we combine the pictures with positive and negative frequencies, and we combine the pictures with positive and negative time axis.

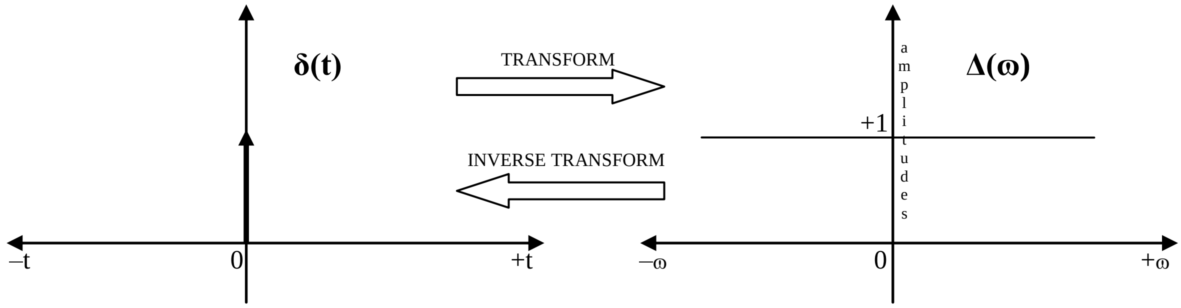




This is known as the SHIFTED UNIT IMPULSE, because it is shifted away from the center/origin in the TIME DOMAIN.

We may now define the UNIT IMPULSE function as located at the origin.

The Unit Impulse Function



$$\delta(t) = \begin{cases} \infty & \text{at } t = 0 \\ 0 & \text{elsewhere} \end{cases}$$

But we must keep in mind that:

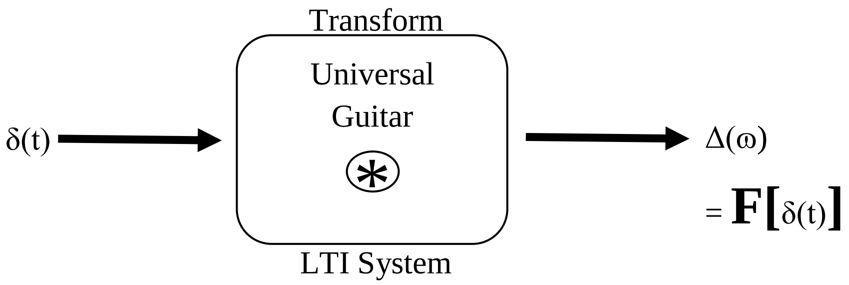
- When the target is on the +x-axis in the SPATIAL DOMAIN, this corresponds to +t instant (t>0) is the TIME DOMAIN and correspondingly positive frequencies are involved in the FREQUENCY DOMAIN.
- When the target is on the –x-axis is the SPATIAL DOMAIN, this corresponds to a –t instant (t<0) in the TIME DOMAIN and correspondingly negative frequencies are involved in the FREQUENCY DOMAIN.
- When the target is at the origin in the SPATIAL DOMAIN then t=0 in the TIME DOMAIN (as depicted above) and all the frequencies (both negative and positive) are involved in the FREQUENCY DOMAIN.

Observe in the SPATIAL DOMAIN at the origin all the frequencies have peak amplitude, whether or not they are rational multiples of each other. Hence for the Unit Impulse (at the origin in the TIME DOMAIN) the spectrum is the entire/***complete*** set of frequencies in the FREQUENCY DOMAIN. And so it is a ***continuous*** line.

The Unit Impulse in the TIME DOMAIN may be depicted/represented by a CONTINUOUS/COMPLETE set of frequencies in the FREQUENCY DOMAIN. Diagrams in DSP text books do not exactly represent the calculations and scenario. They are merely visual aids. Also the variable t in the different functions is heavily over used. It is difficult to tell the context. So we use superscripts to try to make the context more clear.

What is the TRANSFORM of the Unit Impulse function?

Refer to page82 (http://www.nitte.ac.in/nmamit/articles.php?linkId=131&parentId=20&mainId=20&facId=131#). We can visualize



Why do we need the Unit Impulse function?

We want to capture the value/magnitude of a function $f(t)$ at some chosen instant t* and the frequencies present at that instant t*. Every instant t* has a DENSE set of frequencies (that are rational multiples of each other) each with amplitude +1.

To get the frequencies involved we know that CONVOLUTION (with a complete set of carriers) is required. And so we need:

$$f(t) * \textit{something}\,(t^*)$$

This *something* (t*) = $\delta(t^0 - t^*)$,

where the superscript 0 in t^0 is to remind us that at t = 0 we have $\delta(t) = +\infty$.

Because we CONVOLVED $f(t)$ with $\delta(t^0 - t^*)$ and because the Unit Impulse function has all the frequencies present in it(we said earlier that it acts like a Universal Guitar), we get all the frequencies present in $f(t^*)$ at instant t* scaled by the magnitude of $f(t^*)$. This is the IMPULSE RESPONSE. And this information is still in the TIME DOMAIN.

$$\text{IMPULSE RESPONSE } f(t^*) = f(t) * \delta(t^0 - t^*)$$

≡ DENSE set of frequencies present in instant t* with each of the amplitudes (we know is +1) scaled by the magnitude $f(t^*)$.

The transform of function f(t) at chosen instant t* (that is to say the Impulse Response of f(t) at t*) is the transform of f(t*).

FREQUENCY RESPONSE = TRANSFORM [Impulse Response] = $\mathbf{F}\left[f(t^*)\right]$

The FREQUENCY RESPONSE information is in the FREQUENCY DOMAIN

Why not MULTIPLICATION?

To capture the value of any function at some chosen instant t• we may define a new function

$$\delta\bullet\,(t) = \begin{cases} 1 \text{ at } t = 0 \\ 0 \text{ elsewhere} \end{cases}$$

To make it more clear me may denote this by: $\delta\bullet(t^0)$

Then SHIFT $\delta\bullet(t)$ to chosen instant t• to get $\delta\bullet\,(t^0 - t\bullet)$

Then: $f(t\bullet) = f(t)^{\bullet}\,\delta\bullet\,(t^0 - t\bullet)$

Simple multiplication is enough.

f(t•) is a single value , say +5. We may treat it as a simple constant as used in scaling. OR we may treat it as the constant function +5 over (-∞ , +∞). In both cases there is no frequency information - the frequencies involved in f(t) at the instant t•. We want a calculation that will give us both the magnitude of f(t) and the frequencies involved in f(t) at instant t•.

Let us now continue with: IMPULSE RESPONSE $\boldsymbol{f(t^*) = f(t) * \delta(t^0 - t^*)}$

We know that,

$$f(t) * \delta(t^0) = \int_{-\infty}^{+\infty} f(\tau).\,\delta(t^o - \tau).\,d\tau$$
$$= f(t^o)$$

$$f(t) * \delta(t^0 - t^*) = \int_{-\infty}^{+\infty} f(\tau).\,\delta\big(t^o - t^* - \tau\big).\,d\tau$$
$$= \int_{-\infty}^{+\infty} f(\tau).\,\delta\big(\tau - (t^o - t^*)\big).\,d\tau$$
$$= f(t^0 - t^*)$$

Thus we can capture both magnitude and frequency information of $f(t)$ at any chosen instant t^* by CONVOLUTION with $\delta(t - t^*)$ rather than the MULTIPLICATION with the function $\delta\bullet(t)$.

To capture $f(t)$ at my chosen instant t^* AND transform it we need

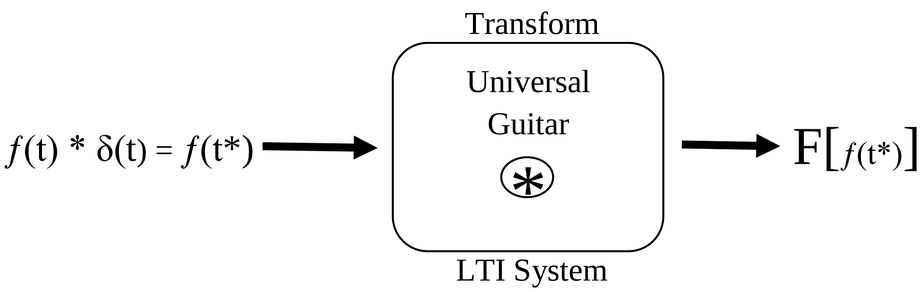
$$\mathrm{F}\left[f(t) * \delta(t^0 - t^*)\right]$$

$$= \int_{-\infty}^{+\infty} \left[f(t) * \delta(t^o - t^*)\right].\, e^{-i2\pi\omega t}\, dt$$

$$= \mathrm{F}\left[f(t^0 - t^*)\right]$$

Even though $f(t) * \delta(t^0 - t^*)$ looks like a single value, it is very different from

1. a simple constant say +5. A simple constant does not have frequencies in it. It is used for scaling.
2. constant function, say $f(t) = +5$ over $[-\infty, +\infty]$. Again there is no frequency information.



Because $f(t)$ is ***aperiodic*** and we want to find $f(t^*)$, that is to say the magnitude of $f(t)$ at ***instant*** t^* and the frequency components present in ***instant*** t^*, we need to perform the CONVOLUTION: $f(t) * \delta(t)$ rather than the CIRCULAR CONVOLUTION. For ***periodic*** $f(t)$ we have the CT-FS, DT-FS and DFT. For ***aperiodic*** $f(t)$ over an ***interval*** rather than an ***instant*** we have the CT-FT.

$$[f(t) * \delta(t)] \circledast \text{ Universal Guitar } \neq f(t) \circledast \text{ Universal Guitar}$$

$$\Downarrow \qquad\qquad\qquad\qquad\qquad \Downarrow$$

$$\mathrm{F}[f(t^*)] \qquad\qquad\qquad\qquad \mathrm{F}[f(t)]$$

Now let us see a real life example of the use of the Unit Impulse function.

We saw in the TWO GUITARS example (http://www.nitte.ac.in/nmamit/articles.php?linkId=131&parentId=20&mainId=20&facId=131#)

how to elicit a frequency response and we even took a look at filtering. However, there are many situations in life where we cannot elicit a frequency response. For example, take a watermelon resting on a table. Its state of ripeness (saatvic, rajasic, tamasic) is a function of time, say melon(t). Now if we play a guitar (no matter how many strings, i.e. all the frequencies) for as long as we like, there is no way we can get a frequency response from the watermelon.

But there is something we can do to get a response. We may impulse the water melon at chosen instant t* and get an Impulse Response melon(t*). This Impulse Response contains many frequencies. We may now transform this Impulse Response to get

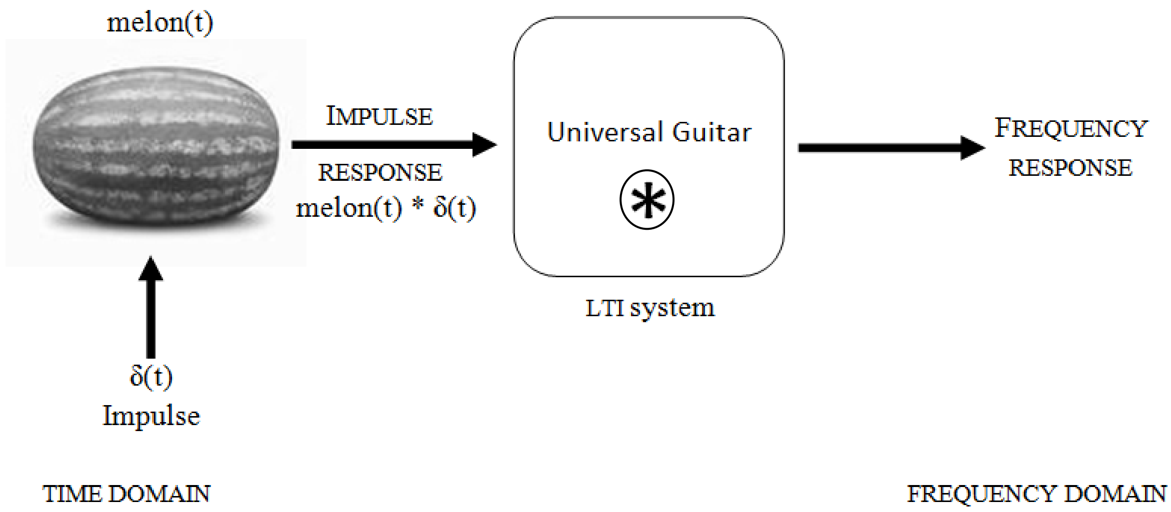
$$F[\text{melon}(t^*)] = F[\text{melon}(t) * \delta(t^*)]$$

Note that we do not need to know melon(t) over some long time interval. We are only interested in melon(t) at some chosen instant t*.

The frequencies we hear /see in the FREQUENCY DOMAIN will tell us the state of watermelon. If there are many lower frequency components with large amplitudes we may say that the watermelon has high water content and hence it is ripe. On the other hand, if the lower frequency components are few and many higher frequency components are present, then we may infer that the water content is low and hence the watermelon is not yet ripe.

Had we chosen another watermelon, we would have got a different set of frequencies with a different magnitude at the same chosen instant t*.

Let us try to understand the Physics.



melon(t) * δ(t – t*) = impulse the watermelon at instant t* to get an Impulse Response.

The Impulse Response melon(t*) contains a set of frequencies which when transformed will tell us something about the state of the watermelon at instant t*.

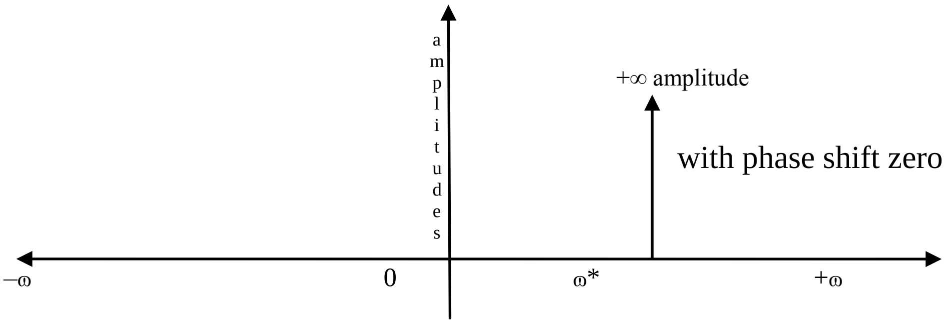
The amplitudes of these frequencies depend on the energy of the Impulse and the damping effect of the rind: the harder the knock/impulse, the larger the amplitudes and hence the louder the sound.

The Frequency Response (transform) of the Impulse Response depends on the medium (LTI System). The Impulse Response of a gong struck (impulse) in air (LTI System = medium) that transforms the Impulse Response will sound (frequency response) different than the same gong struck (impulse) in the same way in helium (LTI System = medium that transforms the Impulse Response).

When it comes to light as the f(t) we may think of the material of the prism as the medium that transforms (LTI System) f(t) at a chosen instant and outputs the spectrum at that instant.

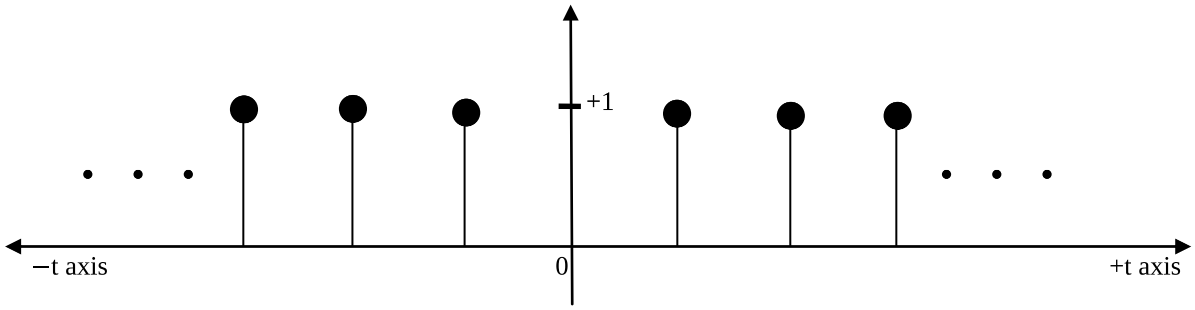
Unit Impulse in FREQUENCY DOMAIN

Let us now look at the Unit Impulse in the FREQUENCY DOMAIN.

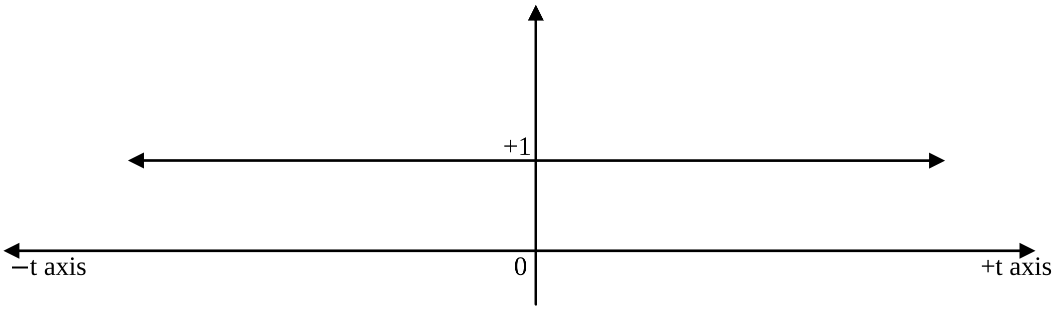


What is the corresponding picture in the TIME DOMAIN?

Looking at the single frequency ϖ*(>0) in the SPATIAL DOMAIN we can see that there are infinitely many peaks with a separation of $2\pi/_{\omega^*}$, the span of a cycle. The target could be on any one of these peaks. Each peak represents a distance in the SPATIAL DOMAIN picture. And each distance in the SPATIAL DOMAIN picture corresponds to an instant (on the +t-axis) in the time domain with the magnitude of +1 due to the +∞ amplitude in the FREQUENCY DOMAIN. So to start with the time domain picture is a discrete set of infinitely many equally spaced time instants (on the +t-axis) each with magnitude +1.



Now depending on the phase (ref. page 4), these peaks (of frequency ϖ^*) can be at any of the continuously many positions within a cycle again with the same separation of $2\pi/_{\omega^*}$, the span of a cycle. So we represent this in the TIME DOMAIN by a continuous line with magnitude +1. Diagrams in DSP text books do not exactly represent the calculations and scenario. They are merely visual aids.

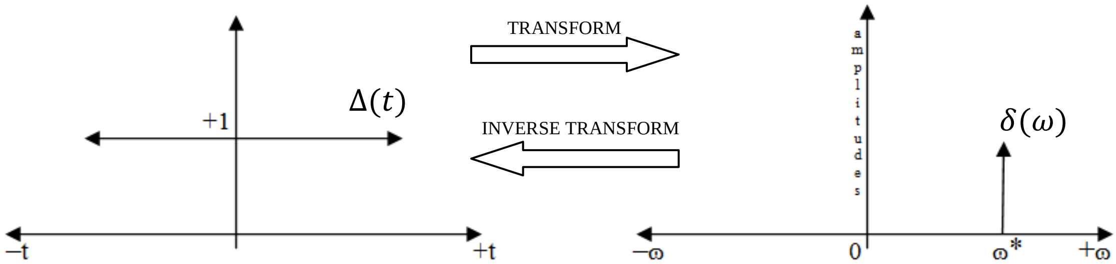


That is to say, given a single frequency ϖ* with the phase ϕ such that $0 \leq \phi < 2\pi$, the target can be anywhere on the time axis.

Since + ϖ * and – ϖ * are complementary, that is to say, they meet/join at the origin to form an eternal sinusoidal, for the sake of graphical representation we may extend the +1 magnitude over the negative time axis. But from calculation point of view we must keep in mind that:

- when ϖ* > 0 this corresponds to instant t* on the +t-axis.
- when ϖ* < 0 this corresponds to instant t* on the –t-axis.
- when ϖ* = 0 this corresponds to t* = 0.

So now we can depict the Unit Impulse in the FREQUENCY DOMAIN as shown in text books.



$\Delta(t)$ and δ(ω) form a transform pair.

Exercise: Explain how you can get the Unit Impulse train in both domains as depicted in the diagrams below. (**Hint**: think in terms of T_d/T_p)