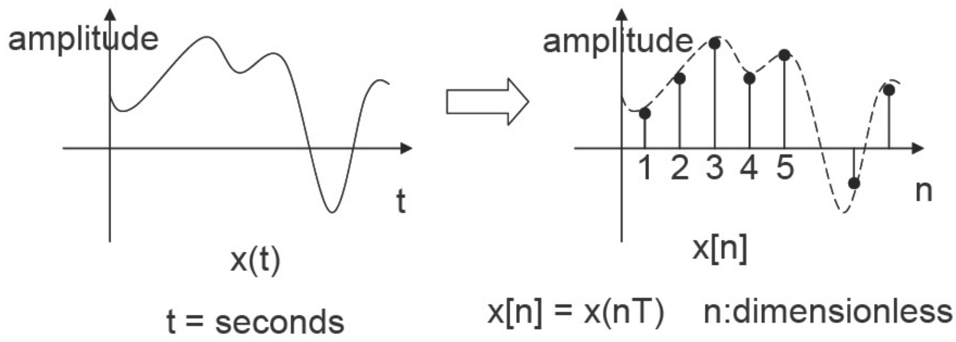
School of Computing
National University of Singapore
CS5240 Theoretical Foundations in Multimedia

1 Why signal processing?

Because we often need to process these:

- Audio — 1D signal $f(t)$
- Image — 2D signal $f(x, y)$
- video — 3D signal $f(x, y, t)$

1.1 Analog-to-Digital conversion

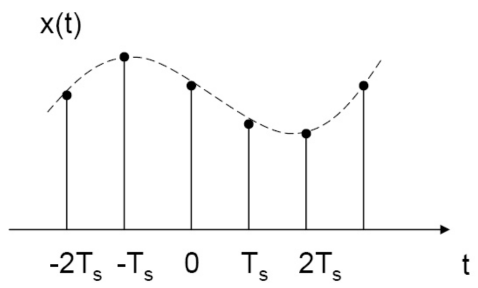


Sampling Period T in seconds, Frequency $\frac{1}{T}$ in Hertz.

Quantization Amplitude x is continuous. Represent using d-bits: 2^d values.

For theoretical considerations, we treat x as continuous. In actual digital computation, x will be quantized. We will study 1D signals first, extend to 2D later.

1.2 Sampling



When converting from continuous $x(t)$ to discrete $x[n]$, we sample periodically.

$$
\begin{aligned}
T_s &= \text{sampling period} \\
f_s &= \frac{1}{T_s} = \text{sampling frequency}
\end{aligned}
$$

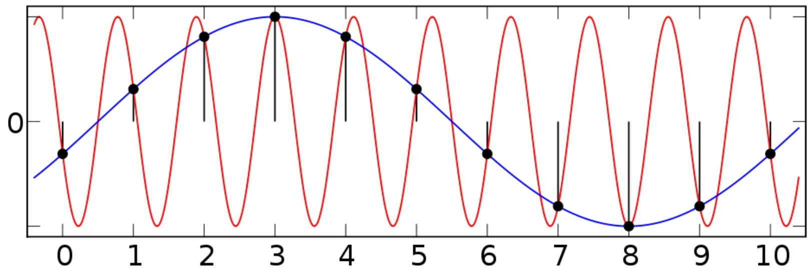


Figure 1: Both the high frequency (red) and low frequency (blue) sinusoids pass through the black samples. Thus using the samples alone, one cannot tell which is the original signal.

Aliasing: Inability to distinguish high frequency signal from a low frequency signal (Figure 1)

Nyquist Sampling Theorem: Let f_H be the highest frequency present in a continuous signal $x(t)$, i.e. $x(t)$ is band-limited. Then the minimum sampling frequency f_s needed to avoid aliasing must be $f_s > 2f_H$.

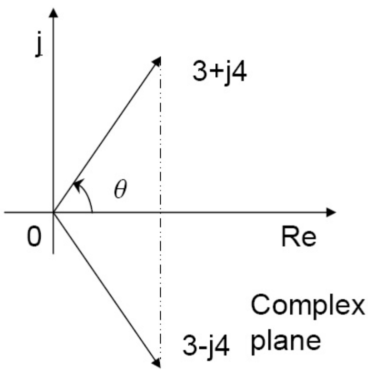
Note: The human ear hears frequencies from 20Hz to 22kHz. That is why music CDs use a sampling frequency $f_s = 44.1$kHz.

For more details on sampling, see Chapter 3 of [1].

2 Review of Complex Numbers & Calculus

2.1 Complex numbers and complex matrices

Imaginary number: $j \triangleq \sqrt{-1}$. So $j^2 = -1$, $j^3 = -j$, $j^4 = 1$
Complex number: $z = a + jb$; $a, b \in \mathbb{R}$. $\quad z \in \mathbb{C}$
Complex conjugate: $z^* = a - jb$



Addition: just like vector addition
Multiplication: $(a + jb)(c + jd) = (ac - bd) + j(ad + bc)$
Cartesian form: $z = a + jb$

Polar form: $z = re^{j\theta}$, r:magnitude; θ: phase angle
Euler relation: $e^{j\theta} = \cos\theta + j\sin\theta$
Example: $z = 3 + j4$
magnitude $r =$ "length" of vector $= \sqrt{3^2 + 4^2} = 5$
phase $\theta : \tan\theta = \frac{4}{3} \Rightarrow \theta = \tan^{-1}(4/3) = 0.9273$ radians.
So $z = 5e^{j\theta}$ where $\theta = 0.9273$
$z^* = (re^{j\theta})^* = re^{-j\theta} = r(\cos\theta - j\sin\theta)$
Multiplication in polar form: $z_1 = r_1 e^{j\theta_1}, z_2 = r_2 e^{j\theta_2}, z_1 z_2 = r_1 r_2 e^{j(\theta_1+\theta_2)}$
i.e. magnitudes are multiplied, phases are summed.

Complex Norm: $|z| = \sqrt{z^* z}$

Inner Product: $< u, v >= u^* v$; $u, v \in \mathbb{C}$
Note:

$$\begin{aligned} e^{j\theta} &= \cos\theta + j\sin\theta, \\ e^{-j\theta} &= \cos\theta - j\sin\theta \\ \Rightarrow e^{j\theta} + e^{-j\theta} &= 2\cos\theta \Rightarrow \cos\theta = \frac{1}{2}(e^{j\theta} + e^{-j\theta}) \\ \Rightarrow e^{j\theta} - e^{-j\theta} &= 2j\sin\theta \Rightarrow \sin\theta = \frac{1}{2j}(e^{j\theta} - e^{-j\theta}) \end{aligned}$$

If $z = z^*$ then z is real, ie. imaginary part of z is 0.

Complex matrices: $\mathbf{A} = [a_{ij}]$; $a_{ij} \in \mathbb{C}$

Hermitian: $\mathbf{A}^H = (\mathbf{A}^\top)^* = (\mathbf{A}^*)^\top$.
e.g. $\begin{bmatrix} j & 1+j \\ 2 & 3-2j \end{bmatrix}^H = \begin{bmatrix} -j & 2 \\ 1-j & 3+2j \end{bmatrix}$

$\mathbf{A}$ is symmetric means $\mathbf{A}^\top = \mathbf{A}$; while $\mathbf{A}$ is Hermitian means $\mathbf{A}^H = \mathbf{A}$;
$\mathbf{A}$ is skew-Hermitian means $\mathbf{A}^H = -\mathbf{A}$.

Vector Norm: $||\mathbf{u}|| = \sqrt{\mathbf{u}^H \mathbf{u}}$

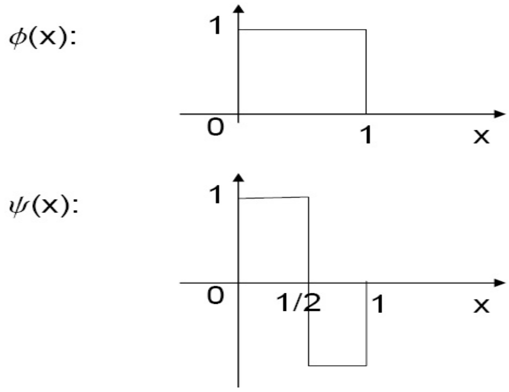
Inner Product: $< \mathbf{u}, \mathbf{v} >= \mathbf{u}^H \mathbf{v} \leftarrow$ complex scalar

Unitary: $\mathbf{A}$ is orthogonal means $\mathbf{A}^\top \mathbf{A} = \mathbf{I}$;
$\mathbf{A}$ is unitary means $\mathbf{A}^H \mathbf{A} = \mathbf{I}$.

2.2 Calculus

We have been dealing with n-dimensional vectors, $\mathbf{x} \in \mathbb{R}^n$ or $\mathbf{x} \in \mathbb{C}$. What if $n \to \infty$? We can think of this as a function.

Let $\phi(x), \psi(x)$ be complex functions of real x, ie. $\phi, \psi : \mathbb{R} \to \mathbb{C}$.
Then, function inner product: $< \phi, \psi >= \int_a^b \phi(x)\psi^*(x)dx$.
Function norm: $||\phi|| = \sqrt{< \phi, \phi >}$. Example:



$$||\phi||^2 =< \phi, \phi > \;=\; \int_0^1 \phi(x)\phi^*(x)dx$$
$$= \int_0^1 1 \cdot dx = [x]_0^1 = 1$$
$$\Rightarrow ||\phi|| = 1$$

$$||\psi||^2 =< \psi, \psi >= \int_0^1 \psi(x)\psi^*(x)dx = \int_0^{1/2} (1)(1)^* dx + \int_{1/2}^1 (-1)(-1)^* = 1$$

$$< \phi, \psi >= \int_0^{1/2} (1)(1)^* dx + \int_{1/2}^1 (1)(-1)^* dx = 0$$

Thus $\phi(x), \psi(x)$ are orthogonal, since $< \phi, \psi >= 0$
In general, $\phi_1(x), \phi_2(x), \cdots \phi_n(x)$ is an orthonormal set of functions if

$$< \phi_i, \phi_k >= \begin{cases} 1 & \text{if } i = k \\ 0 & \text{if } i \neq k \end{cases}$$

Note: These concepts are similar to those in linear algebra.

Fundamental Theorem of Calculus: Let f be a continuous function on the interval $[a, b]$.
Part 1: Then $\frac{d}{dx}\left(\int_a^x f(t)dt\right) = f(x)$
Part 2: $\int_a^b f(x)dx = F(b) - F(a)$ where $F'(x) = f(x)$,
F is called the "anti-derivative" of f.

L'Hopital's Rule: Let $f(c) = g(c) = 0$ for some value c.
Then $\lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{f'(x)}{g'(x)}$
Example: the sinc function (Figure 2): $sinc(x) \triangleq \frac{\sin \pi x}{\pi x}$
What happens when $x = 0$? $sinc(0) = 0/0??$
Answer: use L'Hopital's Rule!

$$\lim_{x\to 0} \frac{\sin \pi x}{\pi x} = \lim_{x\to 0} \frac{\pi \cos \pi x}{\pi} = 1$$

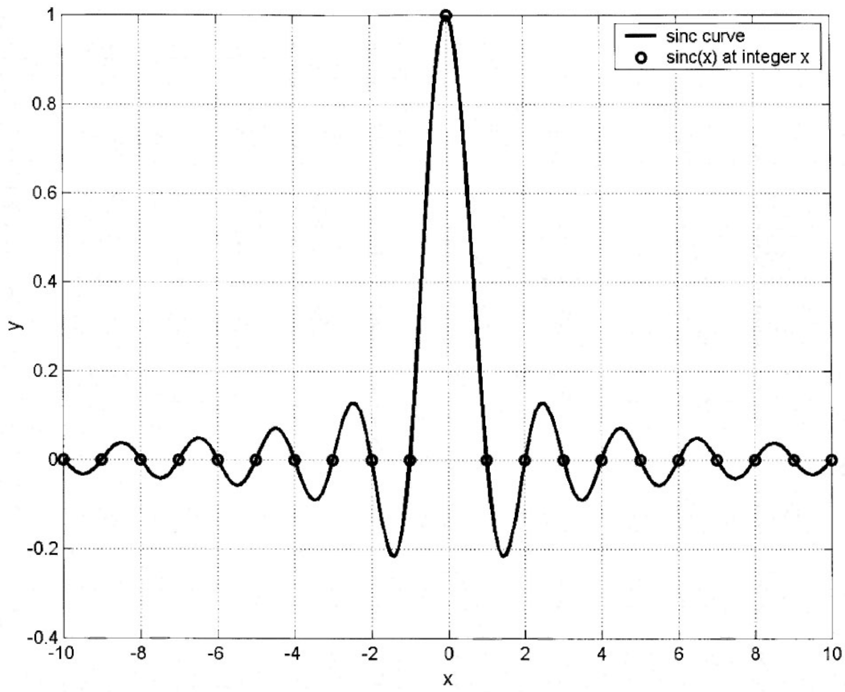


Figure 2: The sinc function. Note that if you sample the sinc function at integer values of x, you get the *impulse* sequence (see next section).

3 Common Sequences

Before we see some common sequences, it would be good to review the following: given a real function $f(x)$, and some constant $k \in \mathbb{R}$, what do we get from:

- $kf(x)$?
- $f(x) + k$?
- $f(x - k)$?
- $f(kx)$?

Figure 3 shows some common sequences (signals). The *unit sample*, or *impulse*, is defined as:

$$\delta[n] = \begin{cases} 1, & \text{if } n = 0, \\ 0, & \text{otherwise} \end{cases}$$

It is zero everywhere except at $n = 0$. The *unit step* is defined as:

$$u[n] = \begin{cases} 1, & \text{if } n \geq 0, \\ 0, & \text{otherwise} \end{cases}$$

The unit step is related to the impulse by:

$$u[n] = \sum_{k=-\infty}^{n} \delta[k]$$

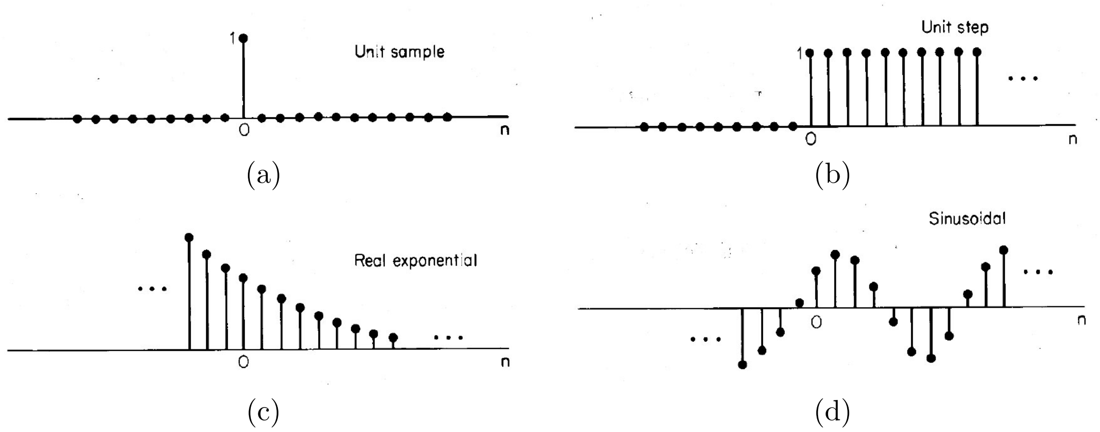


Figure 3: Some common sequences

That is, the unit step at time index n is the accumulation of all impulse values at all previous times. Think of the unit step is the "integral" of the impulse. Alternatively, one may express the impulse as the backward difference of the unit step:

$$\delta[n] = u[n] - u[n-1]$$

Think of the impulse as the "derivative" of the unit step.

An *exponential* sequence is given by: $x[n] = k\alpha^n$. If k and α are real, then the sequence is real. If $0 < \alpha < 1$ and $k > 0$, then the sequence is positive and decreases with increasing n, as in Figure 3(c). If $-1 < \alpha < 0$, the sequence values alternate in sign but its magnitude decreases with increasing n. If $|\alpha| > 1$, the sequence grows in magnitude with n.

Finally, a *sinusoidal* sequence has the form: $x[n] = k\cos(\omega_0 n + \phi)$ for all n, see Figure 3(d). Usually, k is real, and determines the maximum amplitude of the sequence; ω_0 controls the *frequency*, i.e. how fast the sequence oscillates, and ϕ is the *phase* (but do not confuse this with the *phase angle* of a complex number).

One important point to note is that *any sequence* can be expressed as a sum of weighted, shifted impulses. For example, the sequence in Figure 4 may be written as:

$$p[n] = \delta[n+4] + 2\delta[n+2] + 3\delta[n] + (-1)\delta[n-1] + \delta[n-3]$$

In general, any sequence $x[n]$ may be written as:

$$x[n] = \sum_{k=-\infty}^{\infty} x[k]\delta[n-k] \tag{1}$$

We will make use of this important fact later.

4 Systems

Figure 5 shows a general signal processing system, which takes an input sequence $x[n]$ and transforms it into an output signal $y[n]$. Mathematically, we have: $y[n] = T\{x[n]\}$, where

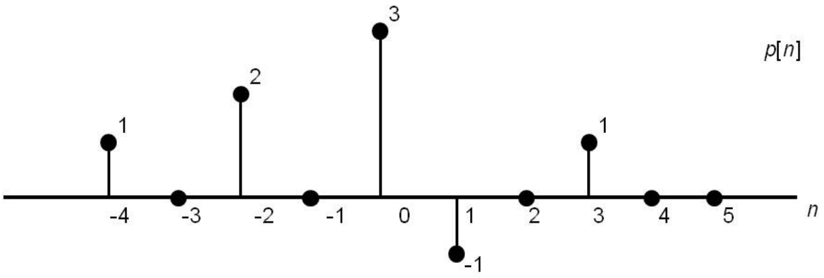


Figure 4: An example sequence.

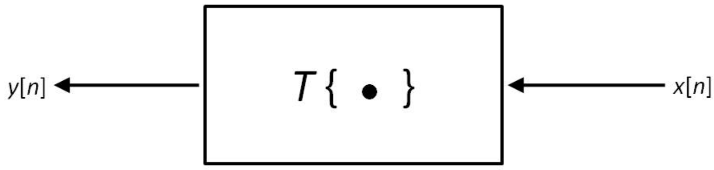


Figure 5: Diagram of a signal processing system.

T is the transformation (or operator) that acts on the input. Of particular interest in signal processing is the class of linear and shift-invariant (or time-invariant) systems. We begin with some definitions.

4.1 Linear Shift-Invariance Systems (Time-Invariant)

Linearity:
Let $x_1[n], x_2[n]$ be two inputs to a system T, and let $y_1[n], y_2[n]$ be the corresponding outputs.

T is said to be *linear* if:

(i) $T\{x_1[n] + x_2[n]\} = T\{x_1[n]\} + T\{x_2[n]\} = y_1[n] + y_2[n]$, and

(ii) For any constant a, $T\{ax[n]\} = aT\{x[n]\} = ay[n]$.

We may combine the above two conditions by saying that for any constants a, b, $T\{ax_1[n] + bx_2[n]\} = ay_1[n] + by_2[n]$. Linearity says that for any linear combination of input signals, the system T produces an output that is the same linear combination of the outputs of the original input signals.

Shift- (Time-) Invariance:

T is said to be *shift-invariant* if for all n_0, $T\{x[n-n_0]\} = y[n-n_0]$.

Shift-invariance says that any shift (or delay) in the input signal results in the output (of the original input) being shifted by the same amount.

Some examples will help clarify these concepts. For each of the systems below, determine if it is linear and/or shift-invariant.

1. The *ideal delay system*: $y[n] = x[n-n_d]$, where n_d is a fixed constant.
2. The *moving average system*:

$$\begin{aligned} y[n] &= \frac{1}{m_1+m_2+1}\sum_{k=-m_1}^{m_2} x[n-k] \\ &= \frac{1}{m_1+m_2+1}\{x[n+m_1]+x[n+m_1-1]+\cdots+x[n]+x[n-1]+\cdots+x[n-m_2]\} \end{aligned}$$

   Note: this computes the average of (m_1+m_2+1) samples around the nth sample.

3. Squaring: $y[n] = (x[n])^2$.
4. The *compressor*: $y[n] = x[Mn]$ for some fixed integer $M > 0$. Note that this discards $(M-1)$ samples out of every M samples, i.e. it outputs every Mth sample.

4.2 Impulse Response

Let $h[n] = T\{\delta[n]\}$ response of T to the impulse input $\delta[n]$. We call $h[n]$ the *impulse response.*

Since T is linear and shift-invariant (LSI),

$$\begin{aligned} y[n] &= T\{x[n]\} \\ &= T\{\sum_{k=-\infty}^{\infty} x[k]\delta[n-k]\} \end{aligned}$$

by using Equation (1). We can push the T into the sum:

$$= \sum_{k=-\infty}^{\infty} x[k]T\{\delta[n-k]\}$$

and finally, $\Rightarrow$ $\boxed{y[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k]}$

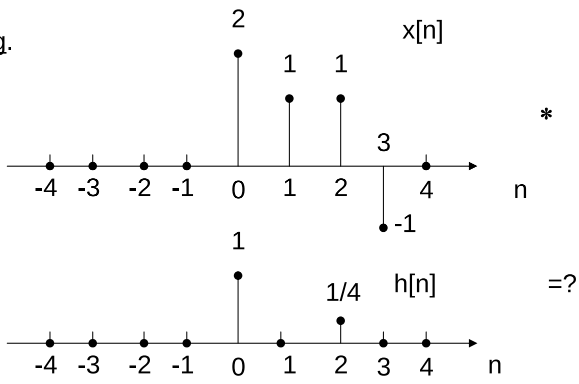
The above is the general response (output) of an LSI system with impulse response $h[n]$ to the input $x[n]$. This operation is so important that we give it a new name:

Convolution sum:

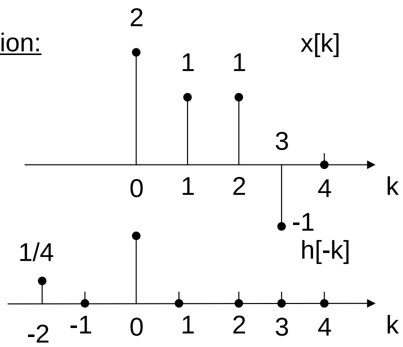
$$x[n] * h[n] \triangleq \sum_{k=-\infty}^{\infty} x[k]h[n-k] \tag{2}$$

Let's see an example:

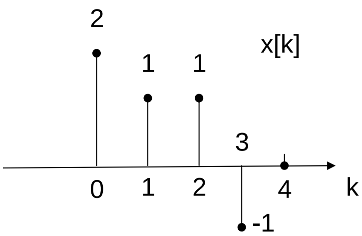
<ins>Eg</ins>.

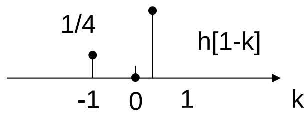


<ins>Solution:</ins>

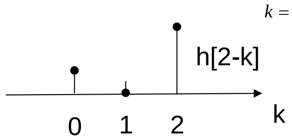


$$y[0] = \sum_{k=-\infty}^{\infty} x[k]h[-k] = 2$$

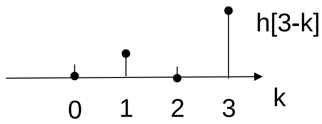




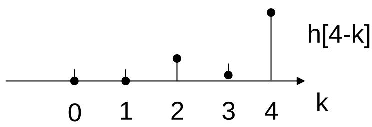
$$y[1] = \sum_{k=-\infty}^{\infty} x[k]h[1-k] = 1$$



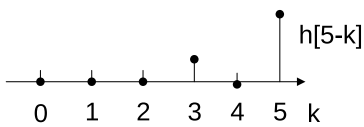
$$y[2] = \sum_{k=-\infty}^{\infty} x[k]h[2-k] = 2\cdot\frac{1}{4} + 1 = \frac{3}{2}$$



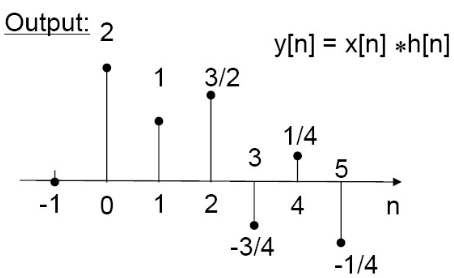
$$y[3] = \sum_{k=-\infty}^{\infty} x[k]h[3-k] = \frac{1}{4} - 1 = -\frac{3}{4}$$



$$y[4] = \frac{1}{4}$$



$$y[5] = -\frac{1}{4}$$



Properties of Convolution:

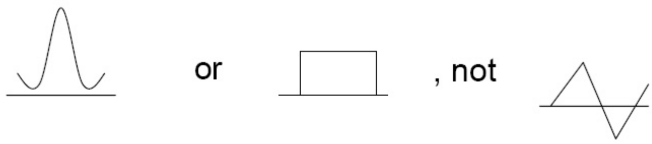
1. $x[n] * \delta[n] = x[n]$.
2. $x[n] * \delta[n-s] = x[n-s]$.
3. Commutative: $x[n] * h[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k]$.
   Change variables: let $m = n-k$, $k = n-m$, which gives
   $\sum_{m=-\infty}^{\infty} x[n-m]h[m] = h[n] * x[n]$.
4. Associative: $(x[n] * h_1[n]) * h_2[n] = x[n] * (h_1[n] * h_2[n])$

5. Distributive: $x[n] * (h_1[n] + h_2[n]) = x[n] * h_1[n] + x[n] * h_2[n]$

Consequence of Central Limit Theorem:

What happens if $x[n]$ is convolved with itself many times? i.e. $x[n] * x[n] * \cdots * x[n]$
Surprise! We always get the same output regardless of $x[n]$!
The output is shaped like a Gaussian, a consequence of the CLT.

Note that $x[n]$ has to be pulse-like:



References

[1] *Discrete-Time Signal Processing* by Alan Oppenheim and Ronald Schafer, Prentice Hall, 1989.