Digital Signal Processing

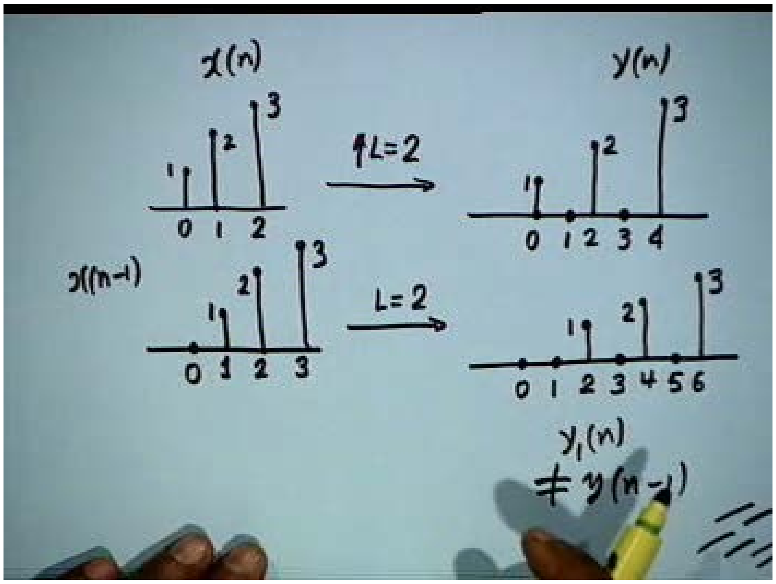
Prof. S. C. Dutta Roy
Department of Electrical Engineering
Indian Institute of Technology, Delhi

Lecture - 5

Digital Systems (contd.), LTI systems, Step and Impulse Responses, Convolution Digital Systems

This is the 5th lecture on DSP and we continue our discussion on digital systems. Today, in particular, we talk about LTI systems i.e. Linear Time Invariant systems, step response of such systems, impulse response, and the relationship between the two, which will ultimately lead to convolution. In the last lecture, we talked about up sampler and also down sampler as examples of digital systems. We also took a moving average system as an example and we introduced the concept of linearity, having two necessary and sufficient conditions, namely the Principle of Homogeneity and the Principle of Superposition. Then we talked about time invariance and time varying. We showed that an up sampler is a time varying system, and so is a down sampler. There was a question about the shape of the waveform in the example of an up-sampler. So I would like to go through this example once more.

(Refer Slide Time: 02:28 – 05:19)

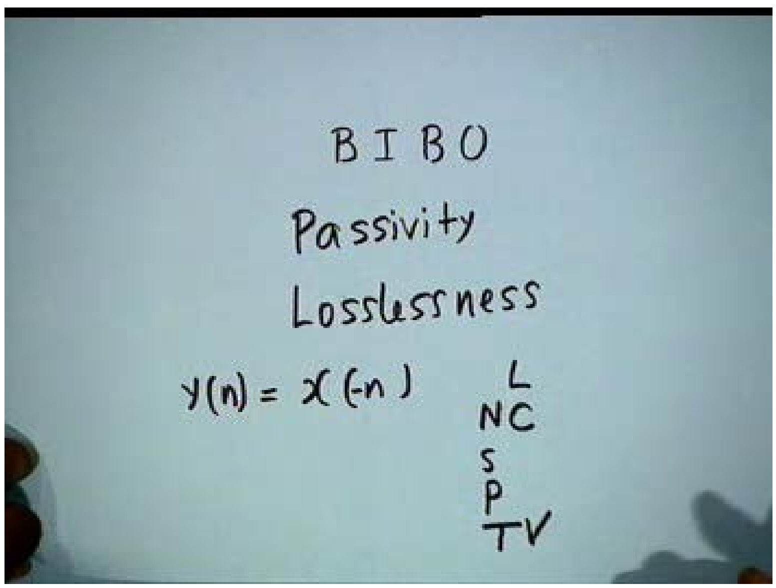


We took x(n) whose amplitude was 1, 2, 3 at n = 0, 1, 2. This has not been drawn to scale but you understand what I mean. When up sampled by L = 2, in y(n), there is a sample at n = 0 of amplitude 1, and then there is a 0 at n = 1, then a sample of height 2 at n = 2, then we have a 0 at n = 3, and then at 4 there is a sample of height 3. Now, if I delay x(n) by one sample, then I start with a 0 then a signal of amplitude 1, next a signal of amplitude 2 and finally a signal of amplitude 3, this is x(n – 1). And when this is up sampled by a factor of 2, what we have is 0, then another 0, then at n = 2 we have a signal of amplitude 1, then we have 0 at n = 3, and then at n = 4, we have a signal of amplitude 2, then a 0 and at n = 6 we have a signal of amplitude 3. It is a little involved concept so you must understand where the trick is. This is y_1(n).

One of the comments was that the waveform that is 1 2 3 with 0s in between is preserved but what about the delay? How many samples are y_1 (n) delayed in comparison to y(n)? The answer is: two samples, whereas x(n – 1) should have led to y(n – 1) if the system was time – invariant. That is, if this was delayed by one sample then we would have concluded that it is a time invariant system. Thus y_1 (n) is not equal to y(n – 1) and therefore it is a time variant system. So waveform preservation is not a sufficient condition for time invariance. Not only waveform preservation, but the delay must be the same as the delay of the input; otherwise it is a time varying system. We also talked about causality and we made a comment about astrologers, which you can ignore. But as far as the electrical engineering systems are concerned, they cannot anticipate what will happen in future.

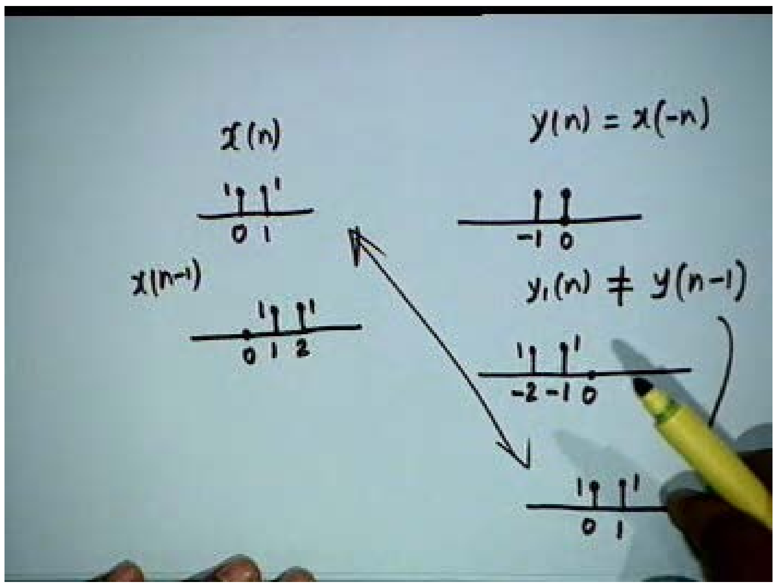
However, non causal DSP can be used in the case of recorded data. That is, the whole data is available and therefore wherever you fix your n = 0, future samples are also available. Non causal DSP is indeed used in several practical applications; one of the most common examples is Geophysical signal processing.

(Refer Slide Time: 06:49 – 08:55)



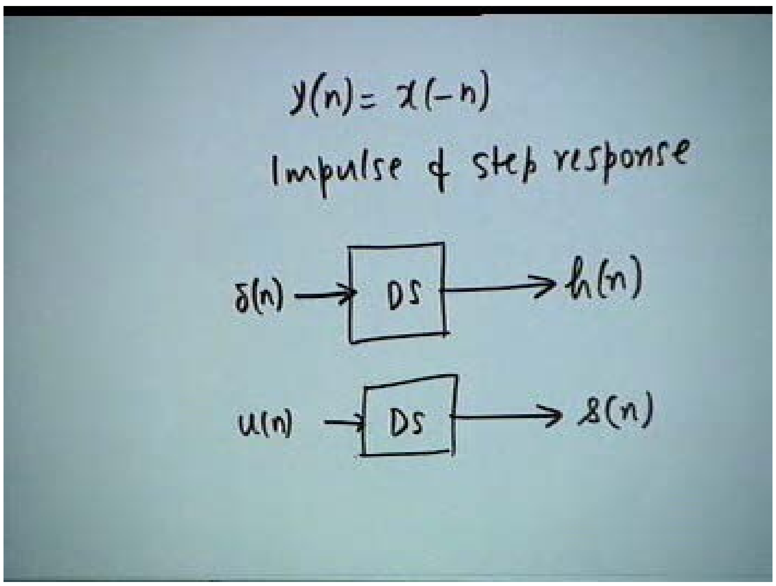
We talked about stability; the kind of stability that we talked about was bounded input and bounded output. In other words, if the input signal is bounded so should be the output signal. If the output signal grows without limit, then it is an unstable system. We introduced the concepts of passivity and losslessness. We also commented that in digital systems, these are artificial concepts, just like energy. Energy of a signal is summation of magnitude squared over all values of n. This is the definition of energy. It is not energy in the sense of analog systems or analog signals. If the output energy is less than or equal to the input energy then we say the system is passive. You realize that a passive digital system can be made active just by multiplying by an arbitrary number. We can always increase the energy by scalar multiplication. This is why it is an artificial concept introduced for the purposes of analysis and design. Similarly when the output energy and the input energy are the same, we say the system is lossless. And we ended the last lecture by taking an example of a system which is $y(n) = x(-n)$ and we said that this system is linear, non causal, stable, passive and time varying. We proved the time varying nature analytically.

(Refer Slide Time: 09:09 – 11:00)



Let us take a very simple example. Let our $x(n)$ be simply the signal of amplitude 1 at $n = 0$ and 1 and zero elsewhere. Then $y(n)$ which is $x(-n)$ will have a sample at $n = 0$ and at $n = -1$. So this is $y(n) = x(-n)$. Now let us delay $x(n)$ by one sample; then $x_1(n) = x(n-1)$ will be 0 at $n = 0$, and 1 at $n = 1$ and $n = 2$. So $y_1(n)$, the response to $x_1(n)$, would be 0 at $n = 0$ and then we shall have 1 at $n = -1$, and 1 at $n = -2$. This is $y_1(n)$ which is obviously not equal to $y(n-1)$. What is $y(n-1)$? What does this signal look like? It is simply 1 at $n = 0$ and 1 at $n = 1$. And don't you see that $y(n-1)$ is the same as $x(n)$? This is incidental, it does not have to be, but you notice that $y_1(n)$ is not equal to $y(n-1)$ and therefore the system is time varying. I must also caution you that in many instances I take examples, after analytical proof or without analytical proof to convince you the truth or falsity of a statement. One million examples are not adequate to prove anything. But you can disprove something by just one counter example. When one is asked for a proof, no amount of examples will suffice. You will have to prove it in general whereas for disproving, one count or example is enough. If you want to prove that a system is non linear, then just apply any one of the principles; superposition or homogeneity. The easiest one to do is to use homogeneity, not superposition. For zero input, does it lead to zero output? If it does not, then the system is non linear, you do not have to go further.

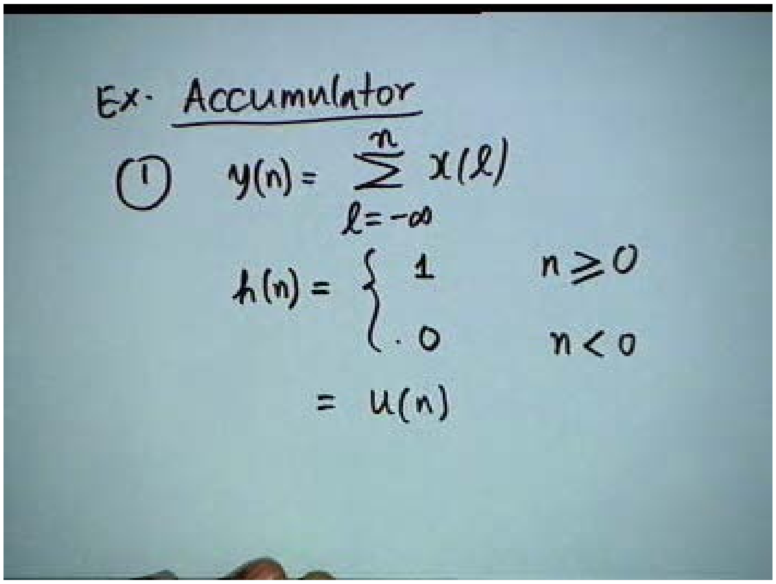
(Refer Slide Time: 12:53 – 14:49)



That $y(n) = x(-n)$ is stable, is obvious because if $x(n)$ is bounded $y(n)$ is bounded by the same quantity.

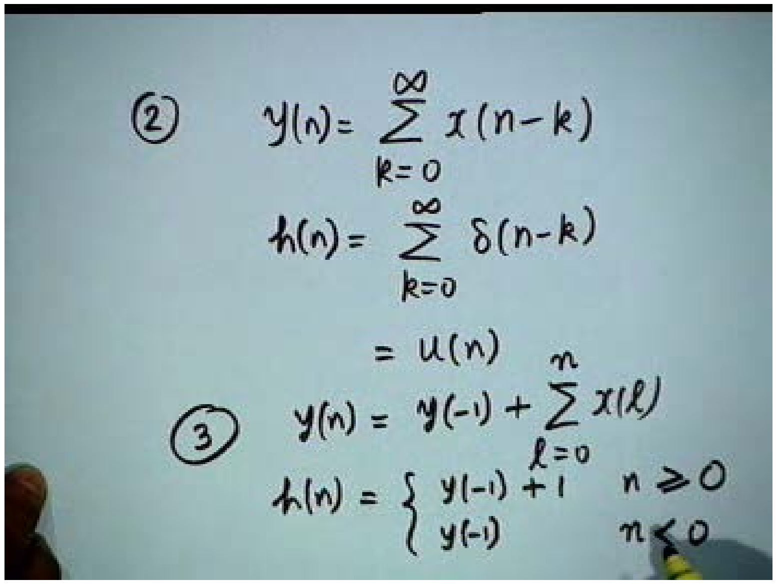
Next we talk about impulse response and step response and the relationship between the two. The definition is very simple. You have a digital system, you apply a $\delta(n)$ to it, whatever the output is we shall denote it by $h(n)$ and call it impulse response. If $h(n)$ is 0 for n less than 0, then the system will be causal because $\delta(n)$ is 0 for $n < 0$. If the system has an output before $n = 0$ it means the system is non causal. Similarly, to the same digital system you apply $u(n)$, the output, usually denoted by $s(n)$, is the unit step response. The symbol s is usually reserved for the analog complex frequency. When there is a chance of confusion we shall not use $s(n)$ to denote unit step response. When we talk about bilinear transformation, transformation from analog to digital or vice versa, we shall use some other symbol. But so long as we do not come to the bridge, we shall not cross it. We shall keep s(n) as the symbol for unit step response.

(Refer Slide Time: 15:02 – 16:28)



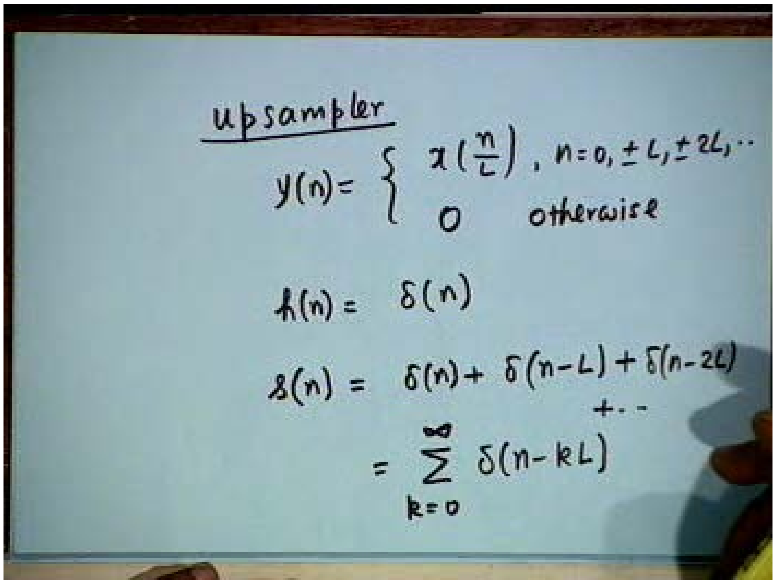
Let us take some examples some of which are a little tricky, although they look very simple. The first example that we take is that of an accumulator. An accumulator can be of various types; ℓet us say that the accumulator is $y(n) = \sum_{l=-\infty}^{n} x(l)$, where n is the present instant of time. Now what do you think its h(n) shall be? To find h(n), we shall have to substitute x(ℓ) by δ(ℓ), which exists only for ℓ = 0 and therefore y(n) is 1 if n is greater than or equal to 0. If n is less than 0, then since δ(ℓ) = 0 for ℓ < 0, we shall have h(n) = 0. Now don't you see that this is precisely the definition of unit step function? And therefore h(n) = u(n).

(Refer Slide Time: 16:29 – 18:36)



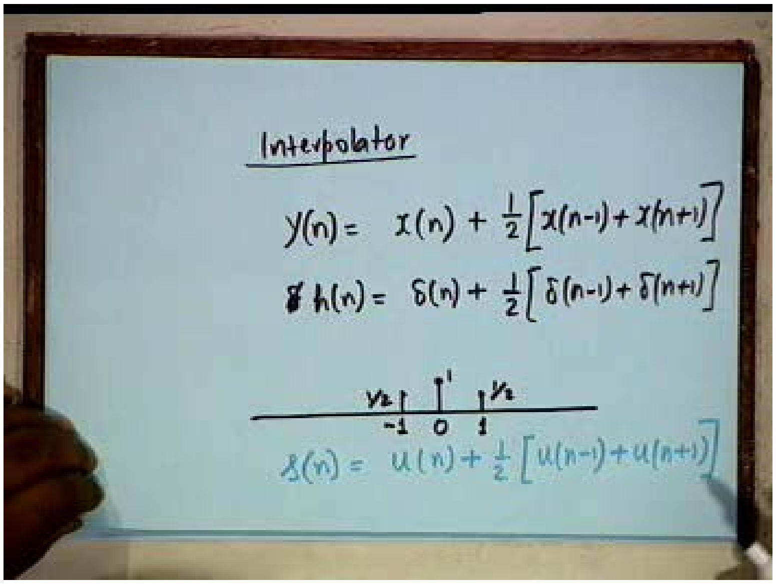
Take another version of the accumulator. Suppose $y(n) = \sum_{k=0}^{\infty} x(n-k)$. Then $h(n) = \sum_{k=0}^{\infty} \delta(n-k)$. It is also precisely u(n). Although the definitions of the two accumulators are different, the impulse response is the same. Suppose we take the third description of an accumulator that is y(n) = y(– 1) + $\sum_{l=0}^{\infty} x(l)$. What do you think its impulse response shall be? The impulse response is h(n) = y(– 1) + 1 if n is greater than or equal to 0. On the other hand, it is simply y(– 1) if n is less than 0. You know that this is a nonlinear system. Non linear systems can also have an impulse response but the other properties that we are going to discuss will not hold for them. Convolution, for example will not hold for a nonlinear system in the form that we are going to present.

(Refer Slide Time: 18:43 – 21:06)



Let us take another example, that of a sampling rate converter. Let us take an up sampler. Up sampler is described by y(n) = x(n/L), in which n = 0 ± L, ± 2L and so on, and it is 0 otherwise. So its h(n) would be simply equal to δ (n). Suppose we want to find the unit step response s(n). You have to substitute x by u, and u (n/ L) = 1 when n = 0 and + L, + 2L etc. Therefore s(n) = δ(n) + δ(n – L) + δ(n – 2L) and so on and in between two samples, it shall have L – 1 0's which amounts to the fact that $s(n) = \sum_{k=0}^{\infty} \delta(n - kL)$. So this is the step response. Similarly you can find out the impulse response of a down sampler.

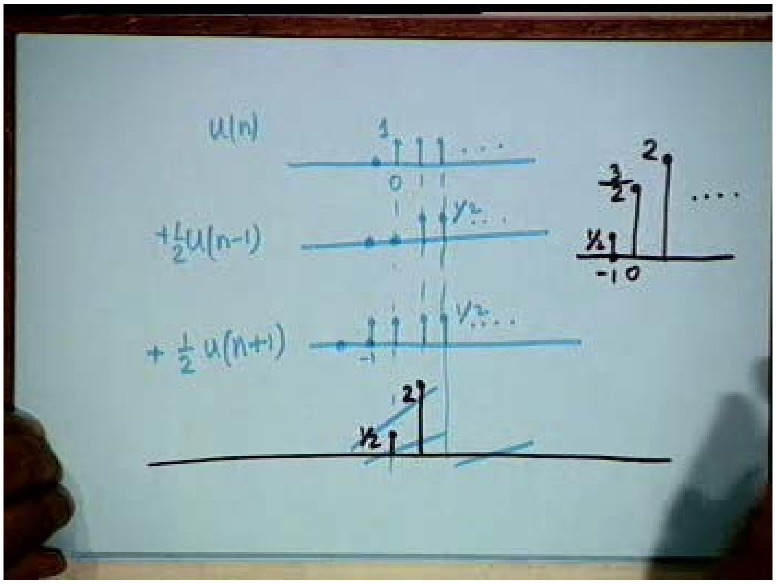
(Refer Slide Time: 21:12 – 22:55)



Suppose we have an interpolator, y(n) = x(n) + (1/2)[(x(n – 1) + x(n + 1)]. Take the present value and to it you add half of the immediate past and the immediate future values. Obviously this is a non causal system because the present output depends on a future input. But what would be its impulse response h(n)? Its impulse response would be δ(n) + ½[(δ(n – 1) + δ(n + 1)]. And if I plot it, it shall have a sample of 1 at n = 0, at n = 1 it shall have a sample of 1/2 and at n = – 1 again, it shall have a sample of amplitude 1/2.

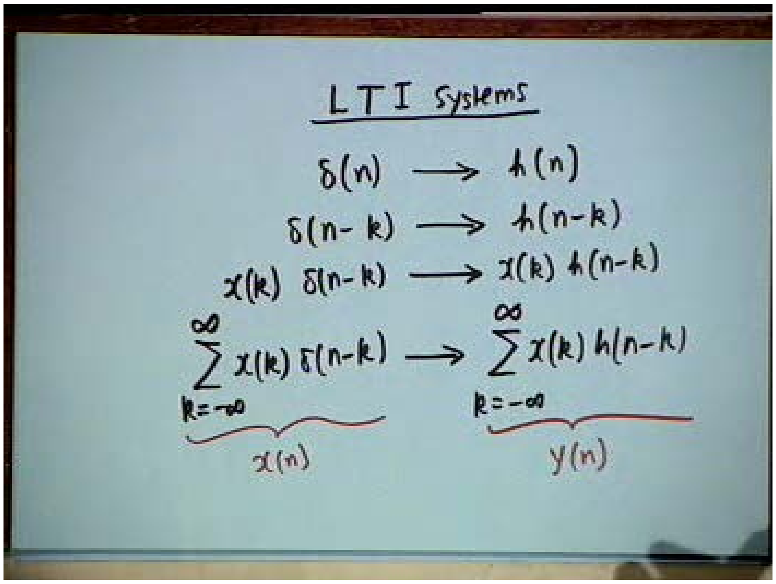
This is what its impulse response looks like. It is interesting to find out what its unit step response would be. The unit step response would be s(n) = u(n) + ½(u (n – 1) + u(n + 1)).

(Refer Slide Time: 22:53 – 25:45)



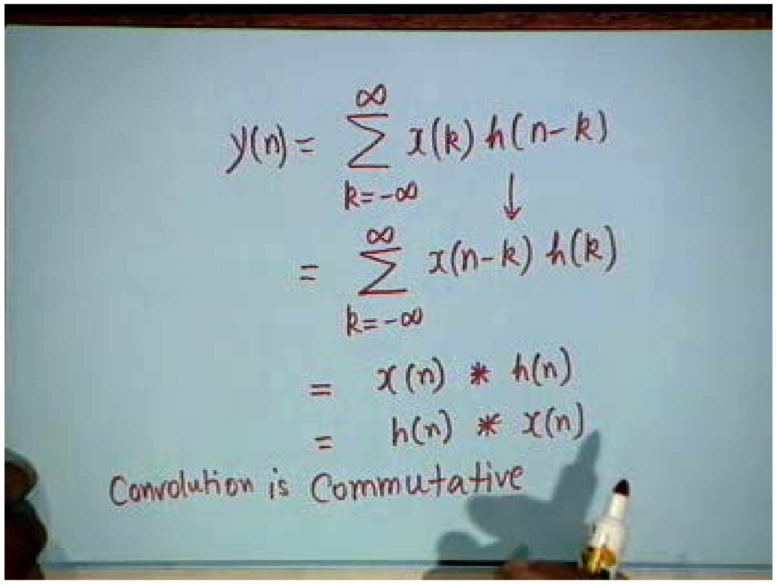
Let us see what this means; u(n) starts at n = 0 and its amplitude is 1 for all values of n. You have to add it to ½ u(n – 1) which means that this starts at n = 1 and amplitude is ½, all through. Then you add to ½ u(n + 1) which starts at n = – 1. Therefore all samples on the left side of n = – 1 are 0, at n = – 1 it would be ½, at n = 0 it is ½ + 1 = 3/2, at n = 1, it is 2 and it remains 2 ever after. So this is how one computes the unit step response.

Refer Slide Time: 26:01 – 27:55)



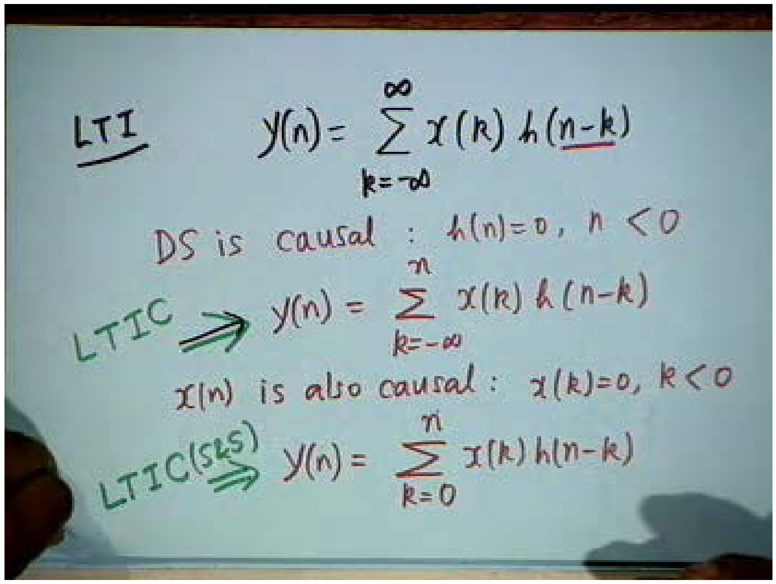
Now we talk about LTI systems; we will not bring non linear systems or time varying systems into our discussion. We talk about purely LTI systems and provide a very simple derivation of convolution. If the system has an impulse response h(n), that is if δ(n) leads to h(n), then δ(n – k), because of time invariance, shall lead to h(n – k). If I multiply δ(n – k) by x(k), a constant, then the output should be multiplied by the same constant, that is, the output shall be x(k) h(n – k); this is an example of application of the principle of homogeneity. If input is multiplied by x(k), output shall also be multiplied by x(k). Then we apply the principle of superposition. That is, we sum this up, we apply all these inputs simultaneously, k going from – infinity to + infinity, in general. Then this should lead to summation[x(k) h(n – k)] where k goes from – infinity to + infinity. But you see the left hand side is precisely x(n); an arbitrary x(n) can always be expressed in this form. The output therefore must be equal to y(n) which is the famous convolution theorem.

(Refer Slide Time: 28:02 – 29:51)



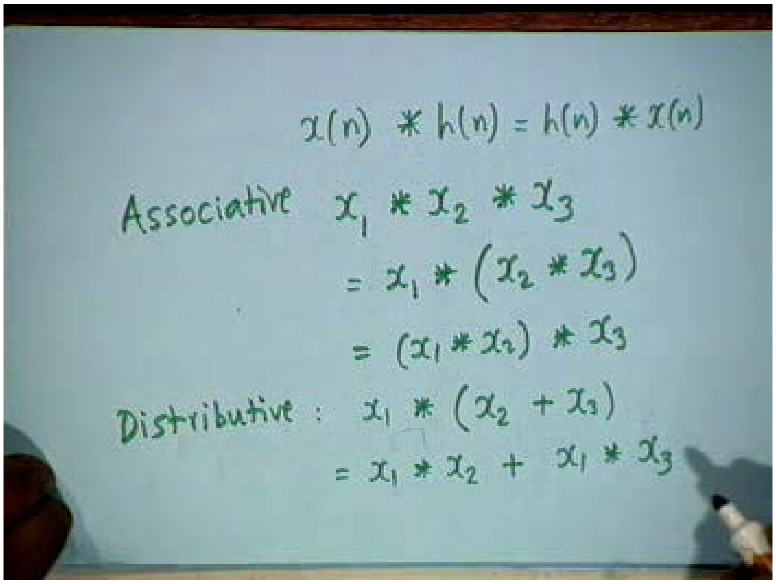
That is, for an LTI system the output y(n) = summation [x(k) h(n – k)] where k goes from – infinity to + infinity. By changing the variable, by putting n – k = r, you can easily show that this is also the same as $\sum_{k=-\infty}^{\infty} x(n-k)$ h (k). This is denoted by y(n) = x(n) * h(n). The star (*) stands for convolution in this particular context. If the star is written above a symbol then it stands for complex conjugate. If the star is written in the horizontal line then it is convolution. And obviously if these two summations are identical then the order of convolution is not important. In other words x(n) * h(n) should be equal to h(n) * x(n). In other words, the operation of convolution is commutative. In general, convolution applies only to Linear Time Invariant systems because you have invoked, in finding the summation, the principles of homogeneity, and superposition, and also the fact that if the input is delayed by a certain number of k samples, then the output is delayed by the same number of samples. That is, we have also invoked time invariance. So this is valid for LTI systems.

(Refer Slide Time: 30:39 – 34:14)



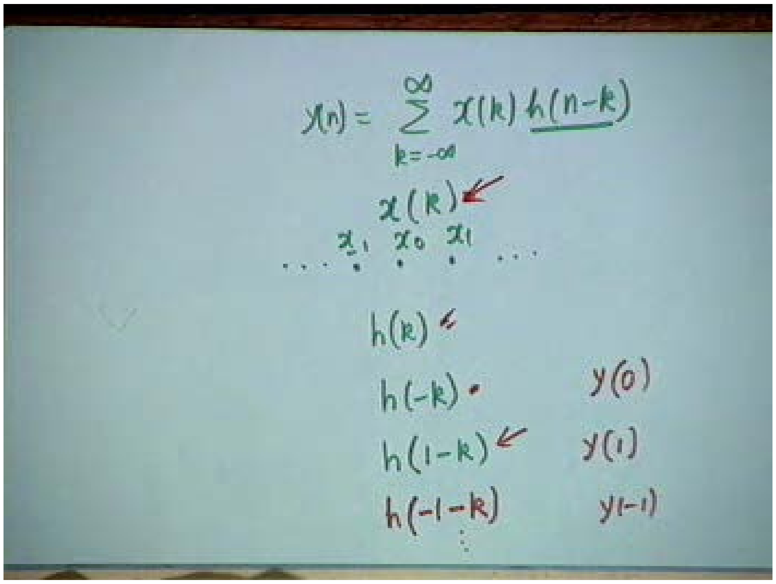
If the system is time varying, then also we can write a convolution summation but it would not be h(n – k); h shall now be a function, to be found out, of n and k and the argument shall not be (n – k). So for a time varying system, instead of h(n – k), we write h(n, k). We shall mostly be concerned with LTI systems. Let us consider some special cases. Suppose the digital system (DS) is causal, then h(n) = 0 for n less than 0. If that is so, then the argument of h(n – k) will become negative when k exceeds n. Therefore the upper limit now should be restricted to n. In the case of a causal system, therefore, the convolution summation simplifies to $\sum_{k=\infty}^{n} x(k)h(n-k)$. Beyond n, the argument of h becomes negative which is 0. Let us say that the input is also causal. That is, x(k) = 0 for k less than 0. If the system is causal and the input is also causal, then our special summation becomes $\sum_{k=0}^{n} x(k)h(n-k)$. In most cases, this is what we shall use. But there is no reason why to a causal system you cannot apply a non causal signal. In the examples that we take, we shall consider non causal signals but a causal system. So y(n) = summation [x(k) h(n – k)], in which k = – infinity to n; this summation is valid for Linear Time Invariant Causal systems LTIC. To LTI, another qualification, C is added to denote that it is also a causal system. For a given situation, you must write the convolution summation correctly, otherwise you will get wrong results.

(Refer Slide Time: 34:23 – 35:55)



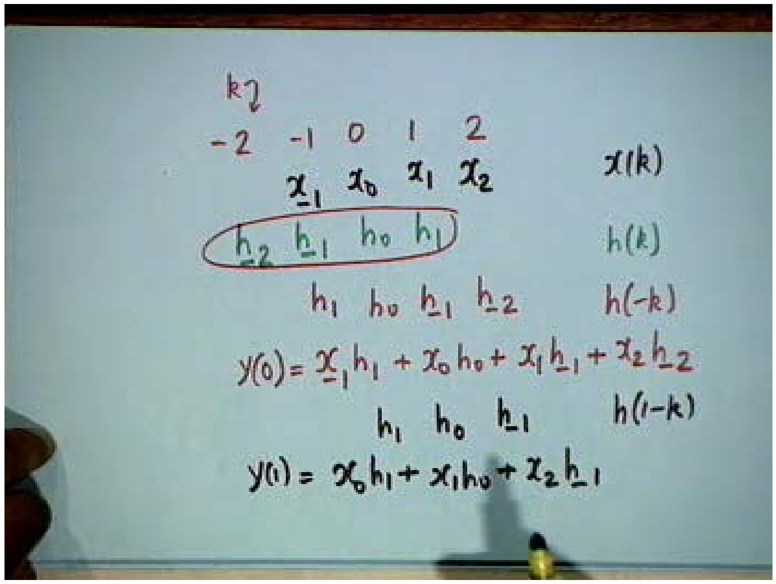
We have already proved that convolution is commutative, that is x(n) * h(n) = h(n) * x(n). By applying the definition, you can very easily show that the convolution operation is also associative. If you have convolution of three signals x_1, x_2 and x_3, then the order in which this is carried out is not important. You can first find the convolution of x_2 and x_3 and then convolve with x_1. Or you can also write (x_1 * x_2) convolved with x_3. This is the Associative property. The other property that the operation of convolution obeys is the Distributive property which says that if you convolve x_1 * (x_2 + x_3) then the result is the same as x_1 * x_2 + x_1 * x_3. The proof, as I said, is extremely simple.

(Refer Slide Time: 36:05 – 38:24)



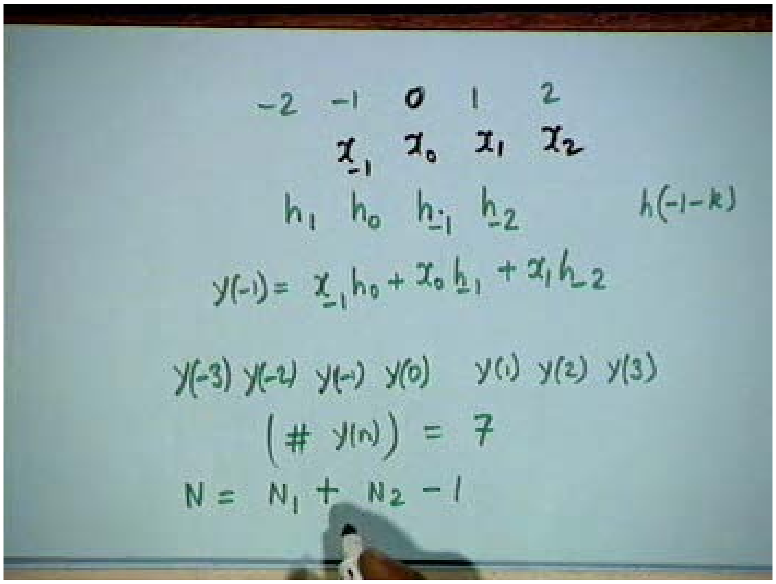
How to compute convolution? This summation $y(n) = \sum_{k=-\infty}^{\infty} [x(k)h(n-k)$ obviously is a very simple operation. But in computing the convolution one has to be careful if you want to do it step by step. I shall give you a trick later. But if you want to do it step by step, then you have to first write x(k) in a row. Then, write h(k) at appropriate places with appropriate values of k below the x(k) signal in the second row. Then find out h(– k) which amounts to taking 0 as the pivot and then flipping it over; you take right side to the left and the left side to the right, then you have h(– k). Next you find y(0) by multiplying the corresponding signals of x(k) and h(– k) and summing them up. Then you find h(1 – k); what is h(1 – k)? It is h(– k) shifted to the right by one sample. And if you multiply the corresponding signals of x(k) and h(1 – k), and sum them up, you get y(1). Similarly you can find out y(– 1) by multiplying x(k) signal by h(– 1 – k), which will mean a shift to the left by one sample and you have to go on computing. Let us take a simple example. Let – 2, – 1, 0, + 1 and + 2 be the set of k over which x(k) and h(k) exist.

(Refer Slide Time: 38:53 – 42:30)



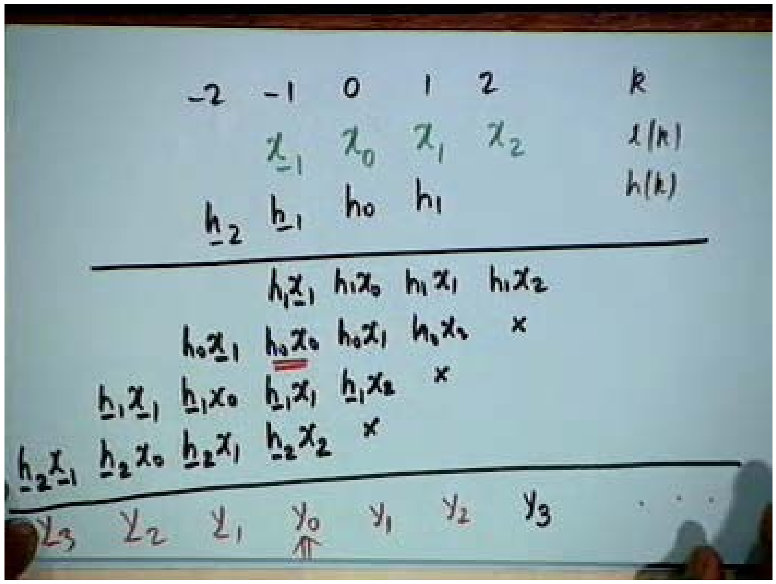
Let $x(k) = \{x_{-1}, x_0, x_1, x_2\}$. To bring variety into experience, let us say $h(k) = \{h_{-2}, h_{-1}, h_0, h_1\}$. I am taking two finite signals. Then $h(-k)$ = will have h_1 and $n = -1$, h_0 at $n = 0$, h_{-1} at $n = 1$ and h_{-2} at $n = 2$. Therefore concentrate now on these corresponding signals. We have $y(0) = x_{-1}h_1$ (the corresponding signals are multiplied) $+ x_0h_0 + x_1h_{-1} + x_2h_{-2}$; there will be four terms. Notice that the sum of the two subscripts is equal to the argument of y; this should always be the case, it is a running check. If you have a term here whose subscript addition does not give the argument of y then you have made a mistake. Our next task would be to find y_1 so you find $h(1 - k)$ by shifting $h(-k)$ by one sample, so I get h_1 at $n = 0$ followed by h_0, h_{-1}, h_{-2}. But I do not write h_{-2}, because the corresponding x(k) signal is 0. So I shift it till there is an overlap with x; I do not shift beyond that. Now I can write y(1) by multiplying the corresponding signals and summing them up. That is $y(1) = x_0h_1 + x_1h_0 + x_2h_{-1}$. Notice that the sum of subscripts of each term is again 1. If I want to construct $y(-1)$ then I shall have to proceed like this; let me write x(k) first, which will be $\{x_{-1}, x_0, x_1, x_2\}$

(Refer Slide Time: 42:45 – 46:49)



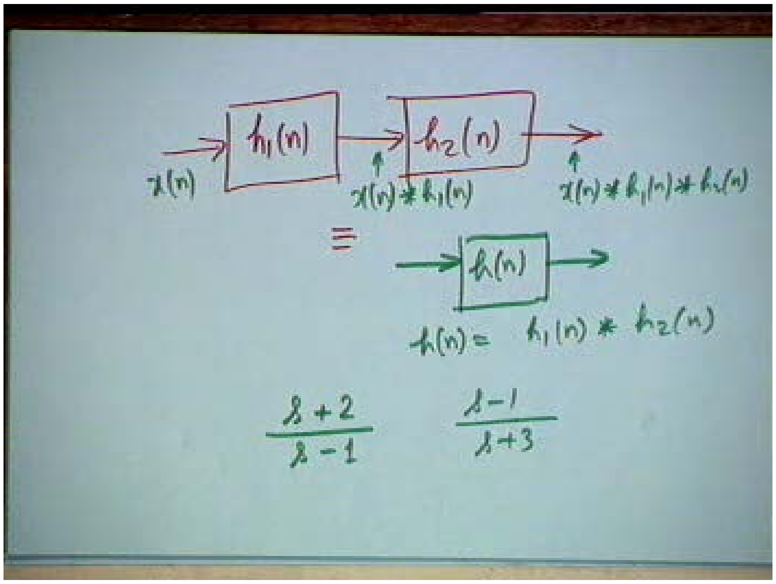
Then I require h(– 1 – k); it would be the same sequence h(– k) shifted one sample to the left. Then at the 0th position it would be $h_{-1,}$ followed by h_{-2} at n = 1; there will be h_0 and n = = – 1, followed by h_1 at n = – 2. So my y(– 1) shall be equal to x_{-1} h_0 + x_0 h_{-1} + $x_1 h_{-2}$; again notice that sum of subscripts in each term is – 1. And you can compute all of them. You will notice that in total there are 7 non-zero samples of y. We shall have y(0) y(1) y(2) y(3). If you go beyond k = 3 then the impulse response goes out of the range of existence of x. Similarly, on the left hand side we can go up to y(– 1), y(– 2) and y(– 3). If we go beyond – 3 to the left, then h goes out of the range of existence of x and there is no overlap. And therefore the number of samples in y(n) is 7 and this is in general true. In general, if the two signals which are being convolved have lengths of N_1 and $N_{2,}$ then y(n) shall have $N_1 + N_2 -$ 1 number of non zero samples. This can be proved rigorously again by writing the signals x(k) and h(– k) in successive rows, and find the shifts needed to the right as well as to the left beyond which there is no overlap between x(k) and h(n – k). Obviously, this is a laborious process and if you a make a mistake, you are done with. All future computations will be defective. Now we show a very simple trick which does not use the graphical method, that is, you do not have to shift and multiply and compute and so on.

(Refer Slide Time: 47:24 – 52:30)



We write the signal x(k) = $\{x_{-1}\ x_0\ x_1\ x_2\}$ and we write h(k) below this at the corresponding positions. This is h(k) = $\{h_{-2}, h_{-1}, h_0, h_1\}$. We now multiply like we multiply two arithmetic numbers. In the two arithmetic numbers we bring them one under the other but here you put them at their appropriate positions (i.e. put them in their places!). And you multiply each of the upper row by each number in the lower row. That is, below the line, you write h_1x_2, h_1x_1, h_1x_0, and then h_1x_{-1} from right (k = 2 position) to left (k = – 1 position). As in arithmetic addition, you leave a column blank and then you take the next number that is h_0. So you put h_0x_2, then h_0x_1, h_0x_0, h_0x_{-1}. Next you leave two columns blank and write $h_{-1}x_2$, $h_{-1}x_1$, $h_{-1}x_0$, $h_{-1}x_{-1}$. The next row would be multiplication by h_{-2}, so you write $h_{-2}x_2$, $h_{-2}x_1$, $h_{-2}x_0$ and $h_{-2}x_{-1}$, after leaving three columns blank and then you add column by column. Please notice the interesting fact that for the first column, the sum of subscripts is 3 and therefore the result of this column should be y_3. Then the sum of the next column shall be y_2 then y_1, followed by y_0. This y_0 is your reference. In general, you shall not been given subscripts; you will be given only numbers so your pivot shall be the column containing $h_0\ x_0$. So this column must give y_0 and once I fix y_0 next ones are y_1, y_2, y_3. To the left of y_0 will come y_{-1}, y_{-2}, y_{-3}. There is nothing else on the right, nothing else on the left, therefore we have got 7 output signals. These are the same as the results obtained by graphical convolution. The logic behind this is extremely simple, you can figure it out. It is simple arithmetic multiplication with a little bit of constraint. There is no carry over here; in arithmetic multiplication, there is a carry over you add it to next term and so on.

(Refer Slide Time: 53:51 – 56:28)



I will just take one more topic. That is, if you interconnect two digital systems having impulse responses $h_1(n)$ and $h_2(n)$ then obviously this is equivalent to a single system of impulse response h(n). If you put an x(n) at input, the output will be $x(n) * h_1(n) * h_2(n)$. So obviously this is equivalent to a single system whose impulse response is h(n), where $h(n) = h_1(n) * h_2(n)$. So in a cascaded system the overall impulse response is the convolution of the individual impulse responses. Now a tricky question: suppose one of the systems is unstable, is it possible that the total system will be stable? Yes, it is possible, as you can see if you consider in terms of poles and zeros. There may be a pole zero cancellation. Take the example of an analog system. I have a system (s + 2)/(s – 1). The other is (s – 1)/(s + 3). In the product, the pole at s = 1 shall be unobservable because it cancels with the zero. So even if the first system tends not to behave, the second system makes it behave. It does not allow the first system to go wild, but in digital systems, there are limitations. For example, you must ensure that there is no saturation in any of the systems. If saturation occurs, then this total system shall be useless.