

Notes on Signal Processing for Linguistics 520

Mark Liberman

October 11, 1998

1 Linearity and Convolution

The aim of this section is clarify the meaning of the phrase: “The effect of any linear, shift-invariant system on an arbitrary input sequence is obtained by convolving the input sequence with the response of the system to a unit impulse.”

To get an idea of what this might be good for, consider some things in the real world that can be successfully modeled as linear shift-invariant systems:

Input	System	Output
laryngeal buzz	vocal tract resonances	vowel sound
noisy recording	noise-rejection filter	less noisy recording
signal	graphic equalizer	modified signal
signal	Dolby encoding	Dolby-encoded signal
Dolby-encoded signal	Dolby decoding	original signal
sound source	room acoustics	reverberant sound

Most of the effort is simply definitional—you have to learn the meaning of technical terms such as “linear,” “convolve,” and so forth. We will also introduce some convenient mathematical notation. Beyond definitions and notation, only some easy high-school-level algebra is required; however, the resulting concept is a powerful one that will enable you to understand quite a bit about digital filtering and speech synthesis.

For mathematical convenience, we treat a digital signal s as an infinitely-long sequence of numbers. We can adapt the mathematical fiction of infinity to everyday finite reality by assuming that all signal values are zero outside of some finite-length sub-sequence. The positions in one of these infinitely-long sequences of numbers are indexed by integers, so that we take $s(n)$ to mean “the n th number in sequence s ,” usually called “ s of n ” for short. Sometimes

we will alternatively use $s(n)$ to refer to the entire sequence s , by thinking of n as a free variable.

We will let an index like n range over negative as well as positive integers, and also zero. Thus

$$s = \{s(n)\}, -\infty < n < \infty,$$

where the curly braces are a notation meaning “set,” so that the whole expression means “the set of numbers $s(n)$ where n takes on all values from minus infinity to infinity.”

We will refer to the individual numbers in a sequence s as *elements* or *samples*. The word *sample* comes from the fact that we usually think of these sequences as discretely-sampled versions of continuous functions, such as the result of sampling an acoustic waveform some finite number of times a second, but in fact nothing that is presented in this section depends on a sequence being anything other than an ordered set of numbers.

The *unit impulse* or *unit sample sequence*, written $\delta(n)$, is a sequence that is one at sample point zero, and zero everywhere else:

$$\delta(n) = \begin{cases} 1 & \text{if } n = 0 \\ 0 & \text{otherwise} \end{cases}$$

The Greek capital sigma, \sum , pronounced *sum*, is used as a notation for adding up a set of numbers, typically by having some variable take on a specified set of values. Thus

$$\sum_{i=1}^5 i$$

is shorthand for

$$1 + 2 + 3 + 4 + 5$$

and

$$\sum_{i=0}^3 f(x_i)$$

is shorthand for

$$f(x_0) + f(x_1) + f(x_2) + f(x_3).$$

The \sum notation is particularly helpful in dealing with sums over *sequences*, in the sense of *sequence* used in this section, as in the following simple example. The *unit step sequence*, written $u(n)$, is a sequence that is zero at all sample points less than zero, and 1 at all sample points greater than or equal to zero:

$$u(n) = \begin{cases} 0 & \text{if } n < 0 \\ 1 & \text{if } n \geq 0 \end{cases}$$

The unit step sequence can also be obtained as a cumulative sum of the unit impulse:

$$u(n) = \sum_{k=-\infty}^n \delta(k)$$

Up to $n = -1$ the sum will be 0, since all the values of $\delta(n)$ for negative n are 0; at $n = 0$ the cumulative sum jumps to 1, since $\delta(0) = 1$; and the cumulative sum stays at 1 for all values of n greater than 0, since all the rest of the values of $\delta(n)$ are 0 again.

This is not a particularly impressive use of the \sum notation, but it should help you to understand that it can be perfectly sensible to talk about infinite sums. Note that we can also express the relationship between $u(n)$ and $\delta(n)$ in the other direction:

$$\delta(n) = u(n) - u(n-1).$$

In general, it is useful to talk about applying the ordinary operations of arithmetic to sequences. Thus we can write the product of sequences x and y as xy , meaning the sequence made up of the products of the corresponding elements:

$$\{x(n)y(n)\}.$$

Likewise the sum of sequences x and y can be written $x + y$, meaning

$$\{x(n) + y(n)\}.$$

A sequence x can be multiplied by a scalar α , with the meaning that each element of x is individually so multiplied:

$$\alpha x = \{\alpha x(n)\}.$$

Finally, a sequence may be shifted by any integer number of sample points:

$$y(n) = x(n - n_0) \text{ for } n_0 \text{ an integer.}$$

We already used this notation when we expressed the unit impulse sequence in terms of the unit step sequence, as the difference between a given sample and the immediately previous sample.

Any sequence can be expressed as a sum of scaled and shifted unit samples. Conceptually, this is trivial: we just make, for each sample of the original sequence, a new sequence whose sole non-zero member is that chosen sample, and we add up all these single-sample sequences to make up the original sequence. Each of these single-sample sequences (really, each sequence contains infinitely many samples, but only one of them is non-zero) can in turn be represented as

a unit impulse (a sample of value 1 located at point 0) scaled by the appropriate value and shifted to the appropriate place. In mathematical language, this is

$$x(n) = \sum_{k=-\infty}^{\infty} x(k)\delta(n-k) \quad (1)$$

where k is a variable that picks out each of the original samples, uses its value to scale the unit impulse, and then shifts the result to the position of the selected sample.

This no doubt seems like a lot of trouble to go to, just to get back the same sequence that we originally started with, but in fact, we will very shortly be able to use equation 1 to perform a marvelous trick, so make sure that you understand it.

A *system* or *transform* T maps an input sequence $x(n)$ onto an output sequence $y(n)$:

$$y(n) = T[x(n)] \quad (2)$$

Thus such a system or transform is a function from sequences to sequences.

Systems or transforms come in a wide variety of types. One important class is known as *linear* systems. We already encountered the concept of linearity in discussing the propagation of sound waves: the linear propagation of sound in air means that the *principle of superposition* applies, so that the pressure disturbance resulting from two sounds propagating through the same region is just the sum of the pressure disturbances corresponding to the individual sounds. Linearity means the same thing as applied to sequences. We can express it mathematically like this:

$$T[\alpha x_1 + \beta x_2] = \alpha T[x_1] + \beta T[x_2] \quad (3)$$

Now we replace the expression $x(n)$ in 2 with the re-expression of $x(n)$ found in 1 to obtain:

$$y(n) = T\left[\sum_{k=-\infty}^{\infty} x(k)\delta(n-k)\right] \quad (4)$$

That is, the sequence consisting of just sample k from x can be rewritten as $x(k)\delta(n-k)$, our old friend the unit impulse scaled to the value of $x(k)$ and shifted to position k . The response of system T to this sequence is just

$$T[x(k)\delta(n-k)],$$

and by linearity we can pull the constant (for this single sample) multiplier $x(k)$ outside of T , giving us

$$x(k)T[\delta(n-k)] \quad (5)$$

as the response of T to a sequence whose only non-zero value is given by $x(k)$.

The original sequence x can be re-expressed as the sum of all of its individual sample values; by linearity, this lets us express $T[x]$ as the sum over all samples of the response of T to a sequence consisting of only the given sample; this response is specified for a particular sample k by equation 5, and so make up the response to the original input, we just need to sum equation 5 over all k , which is:

$$y(n) = \sum_{k=-\infty}^{\infty} x(k)T[\delta(n-k)] \quad (6)$$

Equation 6 has an extraordinary property—it represents the response of system T to an arbitrary input sequence x without applying T to the input x at all! The only thing T operates on is the set of shifted unit impulses, which is independent of x . Having once applied T to the shifted unit impulses, we can calculate $T[x]$ for arbitrary x just by doing the multiplications and additions specified in equation 6.

However, there is one annoyance – we still must calculate $T[\delta(n-k)]$ for every shift k . This would be unnecessary if the response of T to a shifted input was just an output shifted by the same amount. This property, which is called *shift invariance*, holds of many interesting systems. For example, if we put a certain test signal into an acoustic resonator, and then put the same test signal into the same resonator 20 seconds later, we would expect the output to be the same, just shifted in time by the same twenty seconds as the input (as long as the resonator hasn't changed in the meantime).

In mathematical language, a system T is *shift-invariant* if and only if

$$y(n) = T[x(n)] \quad \textbf{implies} \quad y(n-k) = T[x(n-k)]. \quad (7)$$

Thus if we write $h(n)$ for the response of T to the unshifted unit impulse $\delta(n)$, then if T is shift-invariant, the response of T to a unit impulse shifted by k is just $h(n-k)$:

$$h(n) = T[\delta(n)] \quad \textbf{implies} \quad h(n-k) = T[\delta(n-k)]. \quad (8)$$

Thus if T is shift-invariant as well as linear, we can re-write equation 6 as

$$y(n) = \sum_{k=-\infty}^{\infty} x(k)h(n-k). \quad (9)$$

Notice what this means: for any linear shift-invariant system T , once we have calculated its *impulse response* $h(n)$ (its response to a unit impulse at sample point 0), we can forget about T entirely, and just add up scaled and shifted copies of $h(n)$ to calculate the response of T to any input whatsoever. Thus any linear shift-invariant system is completely characterized by its impulse response $h(n)$.

1.1 Convolution

The way of combining two sequences specified by equation 9 is known as *convolution*. For any two sequences x and y , there will be another sequence w obtained by *convolving* x with y , following the equation

$$w(n) = \sum_{k=-\infty}^{\infty} x(n)y(n-k) = x(n) * y(n).$$

The following things are true for convolution in general, as you should be able to verify for yourself by some algebraic manipulation:

$$\begin{array}{lll} x(n) * y(n) = y(n) * x(n) & \text{commutative} & \\ (w(n) * x(n)) * y(n) = w(n) * (x(n) * y(n)) & \text{associative} & \\ x(n) * y(n) + w(n) * y(n) = (x(n) + w(n)) * y(n) & \text{distributive} & \end{array} \quad (10)$$