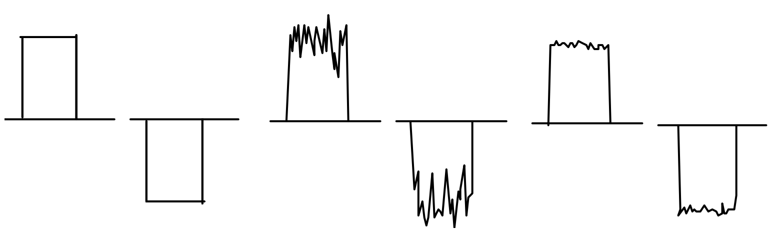
Intuitive Guide to Principles of Communications
www.complextoreal.com

Linear Time Invariant (LTI) Systems and Matched Filter

Matched filter is a theoretical frame work and not the name of a specific type of filter. It is an ideal filter which processes a *received* signal to minimize the effect of noise. Hence, it maximizes the signal to noise ratio (SNR) of the filtered signal.

Let's start by assuming that for a given transmitted signal shape, such as the two square pulses shown below, noise will be added while they are traveling through the medium and also at the receiver, Now we want to filter the received signal to reduce additive noise. What kind of filter should we use? FIR, IIR, Butterworth, what? This is the question we will answer.

It so happens that an optimum filter does exist for each signal shape transmitted and it is a function only of the transmitted pulse shape. Because of its direct relationship to the transmitted pulse shape, it is called a matched filter. At the receiver it filters out the noise such that the received signal SNR is maximum possible and the probability of BER is lowest.



(a) Transmitted signal, square pulses
(b) At the receiver, distorted by a lot of noise.
(c) After the receive filter, looking a good deal more like the transmitted signal.

Fig. 1 – Received vs. transmitted signal shapes.

Let the input to the receiver be $s(t)+n(t)$ where $s(t)$ is the transmitted signal and $n(t)$ the additive noise. Now filter this signal so that additive noise $n(t)$ is removed before demodulation. Since there is no way to remove noise without also removing some of the signal, we want a filter that does this in an optimum way.

Actually what we want is to maximize is the ratio of the received signal to the noise at some particular time, t_d. And not to keep you in any suspense, t_d is at the end of the symbol. Its like waiting until you have all the information before you make a decision, which is a good policy for signal processing as well!

The received signal is designated as $s_o(t) + n_o(t)$. The subscript is there to denote that this is an output signal. The ratio of the signal power to the noise power, or the SNR at the precise instant of the decision time t_d is equal to

$$\eta = SNR = \frac{|s_o(t_d)|^2}{|n_o(t)|^2} \qquad (1)$$

We wish to maximize this number. This is like any other maximization process. We are trying to increase the mean square value of the numerator which is the signal power, and minimize the undesired quantity in the denominator, which is the noise power.

All communications text books spend a whole page or two going through the solution to this maximization, as it is elegant and impressive. The derivation uses Schwartz inequality and other complex terms, is a lot of work but leads to a very simple conclusion. And the conclusion is:

The signal to noise ratio is maximized when the impulse response of that filter is exactly a reversed and time delayed copy of the transmitted signal.

(a) Transmitted signal, (b) the required impulse response of the receive filter.
Fig. 2 – The impulse response of the receive filter should be a reversed copy of the transmitted signal in order to maximize the SNR.

If the signal shape transmitted is 2(a) then the filter that maximizes the SNR has an impulse response shown in this figure 2(b).

What if the transmitted signal is square pulse, then what? Nothing. The signal is symmetrical. The impulse response is the same as itself. This is true of all symmetrical pulses and since nearly all practical pulses are symmetrical, this makes our task easy.

We learned in the BER section that the bit error rate experienced by a signal during demodulation is a function of the signal to noise ratio. A matched filter which maximizes SNR will provide the lowest possible BER. So obviously finding the optimum filter is a worth-while thing to do. The two most common matched filters are Integrate and Dump and Root-raised Cosine.

Linear Time Invariant System

To examine what a matched filter does, we need to visit the concept of a Linear Time Invariant (LTI) system.

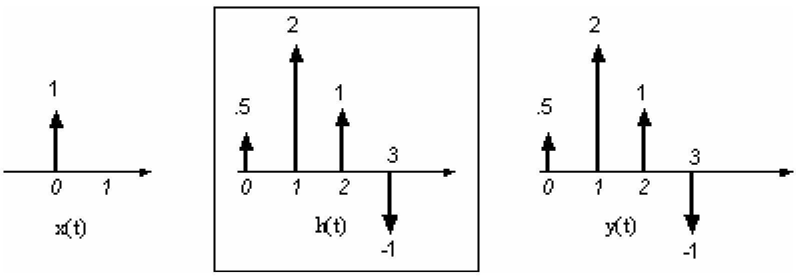


Fig. 3 – The LTI transforms an incoming signal based on a fixed rule.

The middle box in Fig. 3 is the LTI box which does *something* to a signal entering it. In this case, we show a single pulse entering this box. The LTI box takes this pulse, and does *something* to it. All LTI's are different and all do different *something* depending on what you need them to do. All signal processing boxes including filters, amplifiers, etc. can be characterized at LTIs.

The LTI transformation (Eq. 4 and Eq. 5) is the single most important relationship in the field of communications since it gives us a way to relate input and output signals.

The following is model of an LTI.

$$x(t) \xrightarrow[\text{goes in}]{} h(t) \xrightarrow[LTI\ did\ something\ to\ it]{} y(t)$$

x(t) goes in, the LTI does *h(t)* to it and out comes *y(t)* transformed version of *x(t)*. Each pulse goes through the same transformation no matter where in time. Collective response to many pulses spread in time is determined by the Principle of Superposition and forms the transformation of signal through the LTI.

Let's take a look at what happens to the following signal as it goes through this LTI box.

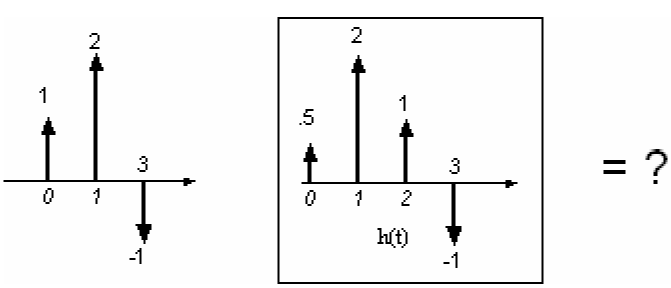


Fig. 4 – Putting a signal composed of many pulses through a LTI

The first pulse to go through the box is at t = 0 time. Its response is the same as the impulse response (because it has a positive unit amplitude!) of the LTI box.

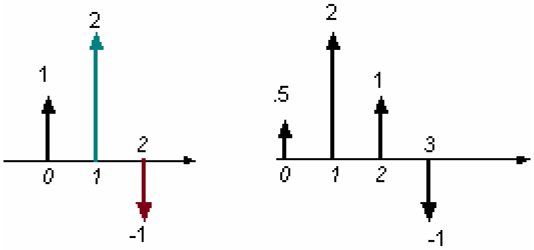


Fig 5a – Response to the impulse at time 0

Now at the next instant, t = 1, the second pulse of magnitude 2 enters the LTI. The response is the same, just twice the magnitude and shifted in time so it starts at t = 1.

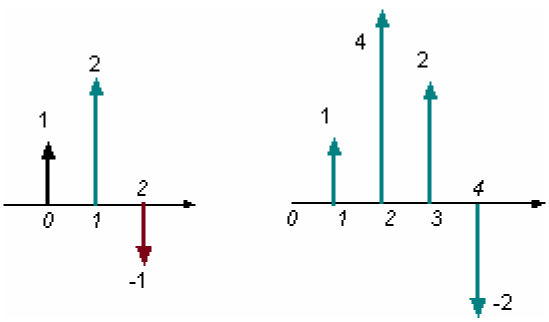


Fig 5b – Response to the impulse at time 1

And similarly the response to the third pulse of magnitude -1 is same as for 1 but reversed and of course starts at t = 2, as shown in Fig. 5c

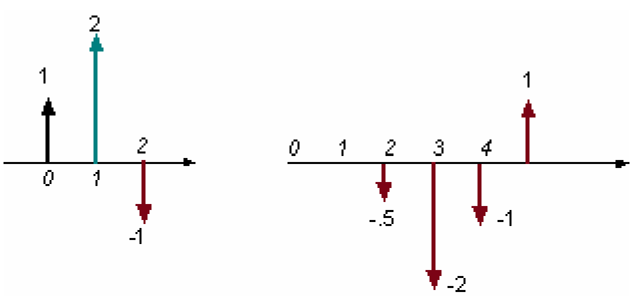


Fig 5c – Response to the impulse at time 2

The signal consists of just 3 pulses. The impulse response of LTI was 4 pulses long for each of the 3 input pulses. The complete response is the sum of all these individual impulse responses. Sadly the output bears no likeness to the input.

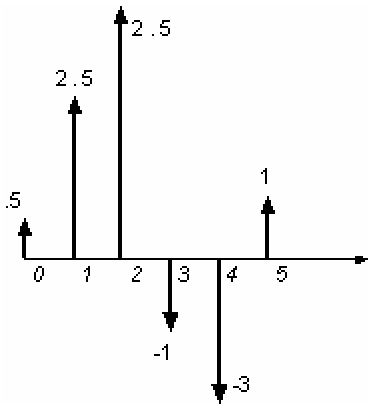


Fig 5d – All responses added up to give the combined response to three pulses.

We can write the output at any one moment in time as

$$y[n] = x[0]h[n] + x[1]h[n-1] + x[2]h[n-2] + \\ x[3]h[n-3] + x[4]h[n-4] + x[5]h[n-5]$$

You notice that in this expression, for each nth value of the response, the input signal x[n] is staying put but the impulse response h[n-k] is moving backwards. We can see this by writing a couple of terms.

$$y[0] = x[0]h[0] + x[1]h[-1] + x[2]h[-2] + \\ x[3]h[-3] + \cancel{x[4]h[-4]} + \cancel{x[5]h[-5]}$$

Since impulse response of the LTI has no 4th or 5th terms, the last two products are crossed out. The next two terms of the output are

$$y[1] = x[0]h[1] + x[1]h[0] + x[2]h[-1] + \\ x[3]h[-2] + x[4]h[-3] + \cancel{x[5]h[-4]}$$

$$y[2] = x[0]h[2] + x[1]h[1] + x[2]h[0] + \\ x[3]h[-1] + x[4]h[-2] + x[5]h[-3]$$

The sum can be written out as the familiar but dreaded convolution equation.

$$y[n] = \sum_{k=-\infty}^{\infty} x[k]\ h[n-k] \qquad (2)$$

(The negative index k in h(n-k) indicates the h(t) shifts in the opposite direction of x(t).) We interpret the response of the LTI as the summation of the weighted and shifted input pulses. It is mathematically a fairly simple concept, although convolution is hard concept to understand intuitively.

The LTI transformation equation is given in time domain by

$$y(t) = h(t) * x(t) \qquad (3)$$

where h(t) is the impulse response of the LTI and y(t) the output/

In frequency domain, we write

$$Y(f) = H(f)\, X(f) \qquad (4)$$

The spectral density of the output signal is given by

$$S_y(f) = S_x(f)\left|H(f)\right|^2 \qquad (5)$$

and another important relationship (both of these given here without proof) about the power in the received signal after it has come out of the LTI is

$$E\left|y^2(t)\right| = \int_{-\infty}^{\infty} \left|H(f)\right|^2 S_x(f)\, df \qquad (6)$$

Convolution and autocorrelation

The *autocorrelation* function (ACF) is similar in concept to convolution, but ACF is a lot easier to grasp.

$$y[n] = \sum_{k=-\infty}^{\infty} x[k]\; x[k-n] \qquad (7)$$

The ACF, as convolution also shifts one of the sequences in time. The only difference between ACF and convolution is that we do not reverse one of the sequences as we do in convolution. Note that the index k here is positive as opposed to Eq. 3. Otherwise the two concepts are identical.

Now the math: You can skip this derivation, if you are feeling math-shy. This part is not essential to understanding what a matched filter is. It is an amazingly simple concept but is wrapped in the mathematical clothes so thick that we can lose sight of what it means.

Statement

The matched filter maximizes the SNR of the filtered signal and has an impulse response that is reverse time-shifted version of the input signal.

Proof

A signal s(t) is transmitted.. It goes through a filter with an impulse response of h(t), It picks up noise n(t) along the way. The output signal is

$$s(t) \quad \rightarrow\rightarrow \quad s_o(t) + n(t)$$

Let's say that the transformation occurs inside a LTI which we will call the matched filter The requirement is that the output signal be as large as possible compared to noise at decision time t = t_d. We can write the expression for average SNR as

$$\eta = SNR = \frac{\left|s_o(t_d)\right|^2}{\left|n_o(t)\right|^2} \qquad (8)$$

Since the matched filter is characterized by its impulse response, we want to find an expression for h(t) that will maximize this SNR.

We will do the maximization in three parts, first we will write the expression for the numerator, the signal, then secondly we will examine the denominator, noise. Thirdly we will use Schwarz inequality to extract useful information about the SNR and make a statement about the optimality of the filter.

The signal

We model the matched filter as a linear time-invariant system which has an impulse response of h(t) and frequency response of H(f). The output out of the LTI is

$$s_o(t) = h(t) * s(t) = F^{-1}\{H(f)S(f)\} \qquad (9)$$

Let's substitute the expression for the inverse Fourier transform

$$s_o(t) = h(t) * s(t) = \int_{-\infty}^{\infty} S(f)H(f)e^{j2\pi ft}df \qquad (10)$$

At time t_d, this equation is written as

$$s_o(t_d) = \int_{-\infty}^{\infty} S(f)H(f)e^{j2\pi f t_d}df \qquad (11)$$

The output power is proportional to the square of this signal amplitude, so we can write

$$S = \left|s_o(t)\right|^2 = \left|\int_{-\infty}^{\infty} H(\omega)S(\omega)e^{j2\pi\pi\omega t}dt\right|^2 \qquad (12)$$

This is the numerator of the expression for SNR that we are trying to maximize.

The Noise

Now let's turn our attention to noise, the denominator term that we need to minimize. The noise is assumed to be Gaussian and has a noise spectral density of $N_0/2$. The noise is zero mean, and has a flat power spectrum wherein all frequencies are represented and have equal amount of power.

We state that the spectral density of the input noise is equal to

$$S_{ni}(\omega) = \frac{N_0}{2} \;\; watts / Hz \qquad (13)$$

The factor of 2 is there because this is a two-sided spectral density. The assumption of $N_0/2$ gives us correct intuitive results. Over a bandwidth B, we get total power which is equal to N_0B as shown below.

$$\begin{aligned} P_n &= \int_{-B}^{B} S_n(f)\, df \\ &= 2\int_{0}^{B} S_n(f)\, df \qquad (14) \\ &= 2\int_{0}^{B} \frac{N_0}{2}\, df = N_0 B \end{aligned}$$

The spectral density of noise-out (no) after the LTI is equal to (using Eq. 5)

$$S_{n0}(f) = S_{ni}(f)\left|H(f)\right|^2 \qquad (15)$$

Now substitute value for the noise-in (ni) spectral density (Eq. 13), we get

$$S_{no}(\omega) = \frac{N_0}{2}\left|H(f)\right|^2 \qquad (16)$$

Employing Eq. 14, the noise power is equal to

$$\left|n_o(t)\right|^2 = \frac{N_0}{2}\int_{-\infty}^{\infty}\left|H(f)\right|^2 df \qquad (17)$$

The fact that noise is assumed to be Gaussian is an important thing to recognize. A matched filter thus derived works as expected only under those conditions. If we have a fading link or other non-Gaussian phenomena then this filter would not be optimum.

Maximizing the ratio

Now we put together both of numerator and denominator is their full glory. Eq. 12 for the numerator and Eq. 17 for the denominator.

$$\eta = SNR = \frac{|s_o(t_d)|^2}{|n_o(t)|^2} = \frac{\left| \int_{-\infty}^{\infty} H(f)S(f)e^{j2\pi f t_d}\, df \right|^2}{\frac{N_0}{2}\int_{-\infty}^{\infty} |H(f)|^2\, df} = \qquad (18)$$

Here is where *Schwarz inequality* comes in handy. The inequality states the equivalence of H(f) and S(f) by the relationship

$$\left| \int_{-\infty}^{\infty} f_1(x)\, f_2(x)\, dx \right|^2 \le \int_{-\infty}^{\infty} |f_1(x)|^2\, df \int_{-\infty}^{\infty} |f_2(x)|^2\, dx \qquad (10)$$

but this hold true only if and only if

$$f_1(x) = k\ f_2^*(x) \qquad (20)$$

where $f_2^*(x)$ is complex conjugate of the function $f_1(x)$. This condition is really important, because we will use it to formulate the final conclusion about the matched filter.

Now set $f_1(x) = H(f)$

And

$f_2(x) = S(f)e^{j2\pi f t_d}$

we write using Eq, 16,

$$\left| \int_{-\infty}^{\infty} H(f)S(f)e^{j2\pi f t_d}\, df \right|^2 \le \underline{\int_{-\infty}^{\infty} |H(f)|^2\, df \int_{-\infty}^{\infty} |S(f)|^2\, df} \qquad (21)$$

Now substitute the right side of the above expression into the numerator of Eq. 15, we get

$$\frac{|s_o(t_d)|^2}{|n_o(t)|^2} = \frac{\left| \int_{-\infty}^{\infty} H(f)S(f)e^{j2\pi f t_d}\, df \right|^2}{\frac{N_0}{2}\int_{-\infty}^{\infty} |H(f)|^2\, df} \le \frac{\cancel{\int_{-\infty}^{\infty} |H(f)|^2\, df} \int_{-\infty}^{\infty} |S(f)|^2\, df}{\frac{N_0}{2}\cancel{\int_{-\infty}^{\infty} |H(f)|^2}\, df} \qquad (22)$$

Apply above to Eq. 17, we can say that the SNR must be less than or equal to

$$\frac{|s_o(t_d)|^2}{\overline{n_o^{\,2}(t)}} \le \frac{2}{N_0}\int_{-\infty}^{\infty}|S(f)|^2\,df \qquad (23)$$

So the maximum SNR possible is equal to

$$\left.\frac{|s_o(t_d)|^2}{\overline{n_o^{\,2}(t)}}\right|_{Max} = \frac{2}{N_0}\int_{-\infty}^{\infty}|S(f)|^2\,df \qquad (24)$$

Now going back to the condition of Eq. 19, this is true only if the following condition is true.

$$H(f) = k\,S^*(f) \qquad (25)$$

Alternately we can say that this means that the magnitude characteristic of the filter should be such that

$$|H(f)| = |S(f)| \qquad (26)$$

This says it all. The filter should be such that it only filters those frequencies in the signal. How simple is that!

This can be written in time domain as

$$h(t) = k\,s^*(t_d - t) \qquad (27)$$

The k is an arbitrary constant and can be used to represent gains through the channel.

This is the final result of the whole thing. The impulse response of the LTI or the matched filter is equal to a time reversed copy of the input signal. And this particular value gives the maximum SNR out of the filter.

Properties of Matched Filter

The output of the LTI system which we will now call the Matched Filter output (as opposed to its impulse response) is the signal $s_o(t_d)$.

Note that Rayleigh's Theorem tells us that the energy of a signal is equal to

$$E = \int_{-\infty}^{\infty} |S(f)|^2 \, df$$

And the output of the matched filter can be written as

$$s_0(t_d) = \frac{1}{2\pi} \int_{-\infty}^{\infty} |S(\omega)|^2 \, d\omega = E \qquad (28)$$

The mean square noise output out of the matched filter is

$$\overline{n_0^2(t)} = \frac{N_0}{2}$$

Which gives us

$$SNR = \frac{2E}{N_0} \qquad (29)$$

No matter what the shape of the transmitted signal, the output of the Matched filter is the energy of the signal at the instant of the decision.

The output of the Matched filter, which is called **matched filer response**, is a signal which may not look like the transmitted signal but has a value of E at the moment of decision.

When we talk about matched filters, be aware if it is *the impulse response* or the *output response*, because they are two different things. The output response may look nothing like the impulse response as we shall see soon. In the above math, the impulse response is h(t) and the output response is y(t).

The Correlation receiver

The process of deriving the matched Filter does not help much in developing an understanding of what the matched Filter is all about. There is n alternate way to look at the Matched Filter process that is much more intuitive. This method is called the Correlation receiver.

For matched filter now we know that h(t) is equal to

$$h(t) = ks^*(t_d - t) \qquad (30)$$

The output response of the matched filter is given by the convolution equation

$$y(t) = \int_{-\infty}^{\infty} x(\lambda)h(t-\lambda)\,d\lambda \qquad (31)$$

for a single bit period t_d, we can write this equation again by changing the limits of integration,

$$y(t) = \int_{0}^{t_d} x(\lambda)h(t-\lambda)\,d\lambda \qquad (32)$$

now lets take the impulse response h(t) (Eq. 26) and shift it by λ so we can plug in its shifted value in the above equation. Correspondingly the right sight shifts by + λ. Remember that the time shift in h(t) has to be opposite to the time shift in s(t) for a matched filter.

$$h(t-\lambda) = k\,s^*(t_d - t + \lambda) \qquad (33)$$

Substituting into Eq. 29 into Eq. 28, we have

$$y(t) = \frac{2k}{N_0}\int_{0}^{t_d}\left[s(\lambda)+n(\lambda)\right]\,s(t_d - t + \lambda)\,d\lambda$$

Now set t = t_d, and we can do that because it we are only interested in the results of this operation at the precise moment when t = t_d or the end of the bit time when we will make a decision about what was sent.

$$y(t) = \frac{2k}{N_0}\int_{0}^{t_d}\left[s(\lambda)+n(\lambda)\right]\,s(\lambda)\,d\lambda \qquad (34)$$

The above expression is exactly the same as doing cross-correlation between the input signal which includes noise and the original signal without noise. This leads us to an alternative way to implement a matched filter, called the correlator.

The correlation receiver is exactly the same thing mathematically as the matched filter but much easier to comprehend and much easier to implement in hardware.

Examples

In Fig 6, we show the matched filter output for a square pulse shown. Impulse response of a square pulse is itself. Note what the filter output value at the end of the symbol period which is 5. The value is maximum at t = 5, the end of the symbol period. The receiver now makes a decision based on this value and declares the bit to be 1. Then the second half of the output is not needed, so the value (of 5) is discarded from memory, the registers are flushed and the process starts anew with the next symbol.

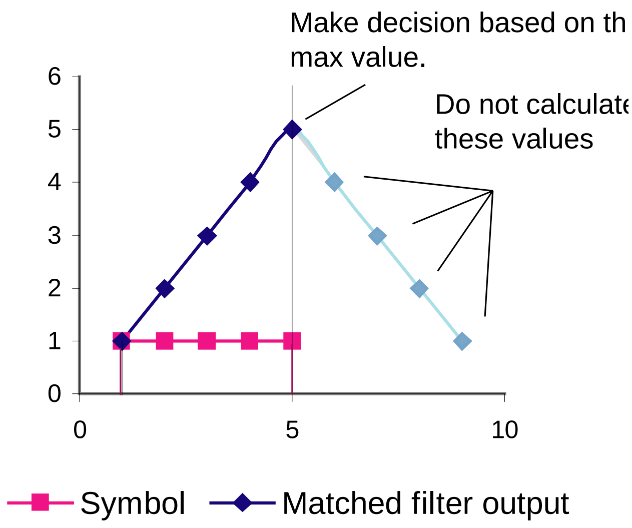


Fig. 6 – A square pulse and its matched filter output.

What's interesting is that the output of the matched filter is maximum at t = 5, the last sample of the symbol. Each point in between amounts to an integration of the symbol up to that point. So although we did convolution, what we got is an integration, exactly the same thing we would get if we do cross-correlation between the transmitted symbol(the square pulse) and the saved reversed replica of this symbol (also a square pulse).

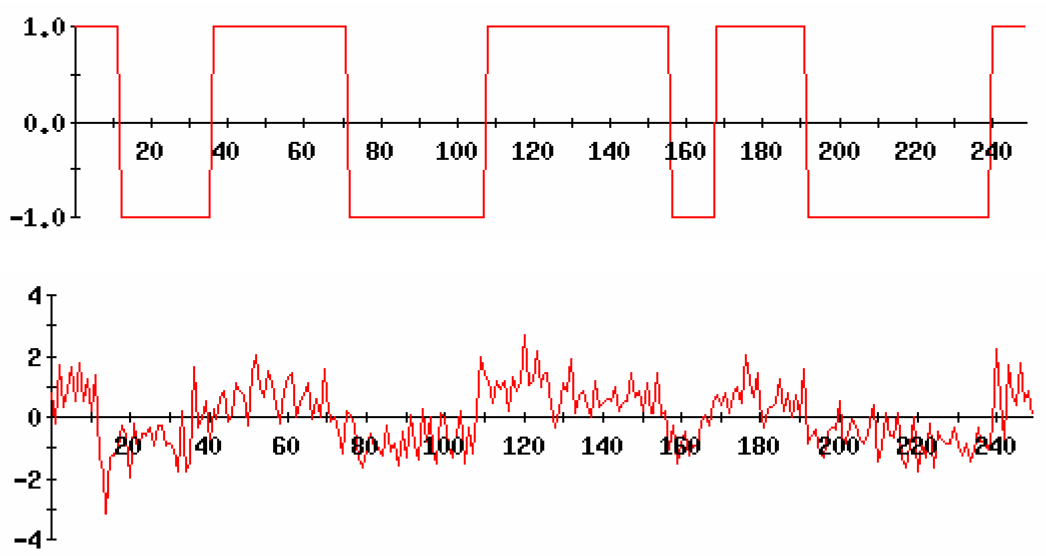


Fig. 7 – Transmitted and received pulses

Fig. 7 shows a square pulse signal and its noisy received counterpart. Fig. 8 shows the timing pulse and the output of the filter. This result gives us the integrate and dump filter. We integrate the received symbol up to the last sample, dump this value to the slicer and the reset to start anew with the next symbol. Depending on the polarity of the next symbol, the next integration may be negative or positive as shown below. A slicer then samples the signal after the dump and makes a bit decision based on a simple logic.

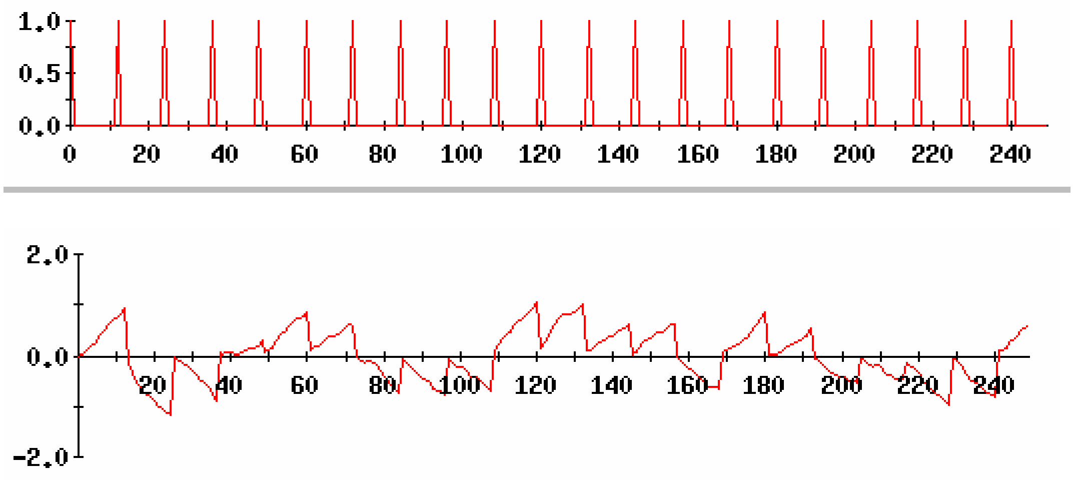


Fig 8 – Integrate and Dump output

Integrate received amplitudes over the symbol width which is 12 samples in this case. At the end of 12 samples, a trigger tells the integrator to stop, dump the values to the slicer and start again. So what we see are little triangles, which is the integration of the square pulse. The perturbation comes from the statistical variation of the added noise value.

Situations where Integrate and Dump filter in not optimum

A triangular pulse shape

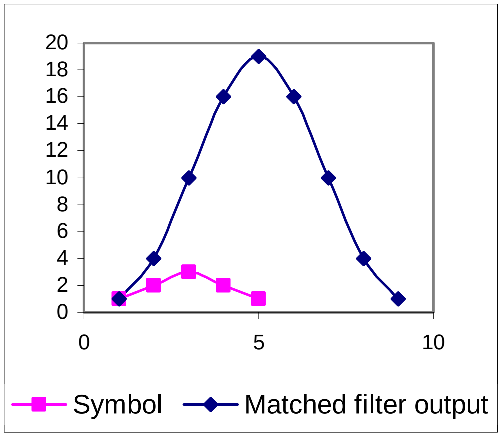


Fig. 9 – A triangular pulse and its matched filter output

In this case, we have a triangular pulse. Note once again, that the out put of the matched filter is the integration of the triangular pulse with maximum occurring at t = 5, a value of 19 which is the area under the pulse. Now once again, we could integrate and make a decision based on the max value but it turns out the integrate and dump filter is not optimum for this case.

Fig. 10 – The effect of noise on the average signal to noise ratio is different for these two pulses.

For a square pulse the signal to noise ratio is constant whereas for the second pulse, the SNR is low at the edges and is maximum in the center for the very obvious reason that the signal is maximum at the center. The integrate and dump does a fine job for the square pulse but it would exaggerate the low SNR region over the high SNR, so is not an optimum filter for anything other than square pulse.

What is optimum for this type of pulse? Since the pulse is symmetrical, the impulse response is the same as the pulse, the output is just the correlation of the two and the best value occurs at the center, hence for this type of signal, we just do a center slice and make a decision based on the center value.

For most pulses other than square, a center sampling is optimum for decoding purposes, because the SNR is maximum at this point.

Implementing Matched filter as a correlator

Let's take the pulse of Fig. 11. The transmitted pulse and the received pulse with noise are shown.

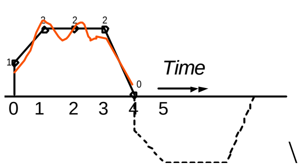


Fig 11 – Transmitted and received pulse

Transmitted pulse	1	2	2	2	0
Received pulse	0.8	2.2	2.2	1.8	0.3

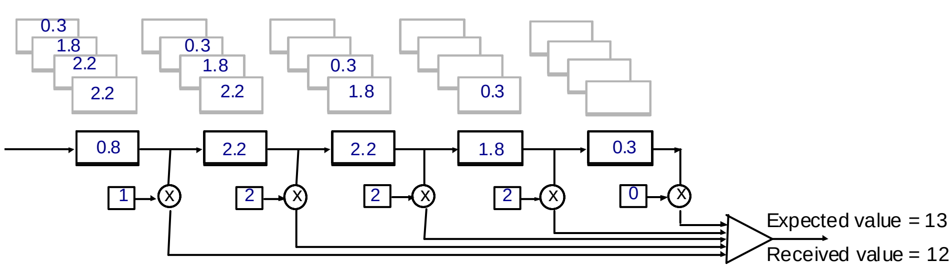


Fig 12 – Matched filter in process (note how similar it looks to a correlator,)

In the receiver we have stored the expected pulse shape in reverse as shown in the boxes next to the multipliers. Now come the samples of the received signal. This is a little confusing because they come in a reverse order, that is 0.8, 2.2, 2.2, 1.8, 0.3., with .8 at first, then 2.2, then 2.2 and so on.

The samples are being fed into the registers. Nothing is happening until all five registers are full. Now perhaps you can see the need for reversing the stored impulse response in the receiver. This is so that the received samples line up properly and we will get a maximum response. Now we do a multiplication of the expected with the received. The expected value is 13. the received value in this case is 13.2.

.8 x 1 + 2 x 2.2 + 2 x 2.2 + 1.8 x 2 + 0 x .3 = 13.2

Based on the rule that if this number is greater than 0, then decide in favor of pulse 1 if less than 1 then zero decide in favor of other pulse.

The output of a matched filer is the energy of the signal

An optimum receiver for AWGN channel does two functions. One is the matched filter which is often the correlator and the other a detector which makes decision about the pulse. The correlator correlates the received signal with two stored pulses, s_0 and s_1 (assuming this is a binary signal). The result is two different output values and these two values are then fed to the detector which makes a decision about what was sent. The output values at time t are given by the expressions

$$\begin{aligned} y_0(t) &= \int_0^t y(\tau)\, \underline{s_0(\tau)}\, d\tau \\ y_1(t) &= \int_0^t y(\tau)\, \underline{s_1(\tau)}\, d\tau \end{aligned} \qquad (35)$$

Example.

A signal y(t) is received. It contains the transmitted signal and noise component. Now we do the two correlations.

$$y(t) = s_0(t) + n(t) \qquad 0 \le t \le t_d$$

$$\begin{aligned} y_0 &= \int_0^{t_d} y(t)\, \underline{s_0(t)}\, dt \\ &= \int_0^{t_d} s_0(t)\, s_0(t)\, dt + \int_0^{t_d} n(t)\, s_0(t)\, dt \qquad (36) \\ &= \quad E + n_0 \end{aligned}$$

Recognize that the first term is the energy of the signal.

$$E = A^2 t_d \qquad (37)$$

which is kind of a autocorrelation of the signal at time t_d. The same thing with the other stored pulse.

$$\begin{aligned} y_1 &= \int_0^{t_d} y(t)\, \underline{s_1(t)}\, dt \\ &= \cancel{\int_0^{t_d} s_0(t)\, s_1(t)\, dt} + \int_0^{t_d} n(t)\, s_1(t)\, dt \qquad (38) \\ &= n_0 \end{aligned}$$

where noise components are given by

$$\begin{aligned} n_0 &= \int_0^{t_d} n(t)\, s_0(t)\, dt \\ n_1 &= \int_0^{t_d} n(t)\, s_1(t)\, dt \end{aligned} \qquad (39)$$

The first term is zero because the signals are assumed to be orthogonal. So we get two results, y_1 and y_0, one is the energy plus noise and the other just noise, so assuming that noise power is less than signal energy, we would make the decision in favor of the higher level.

A root raised cosine-like pulse

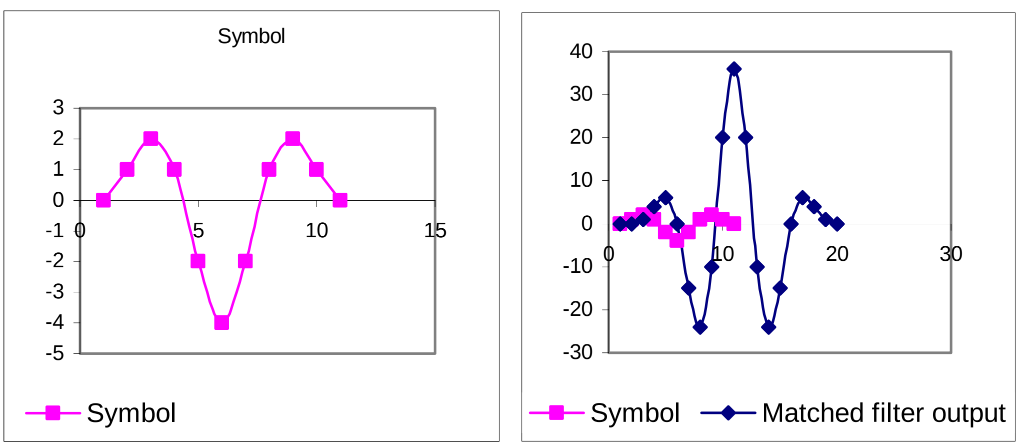
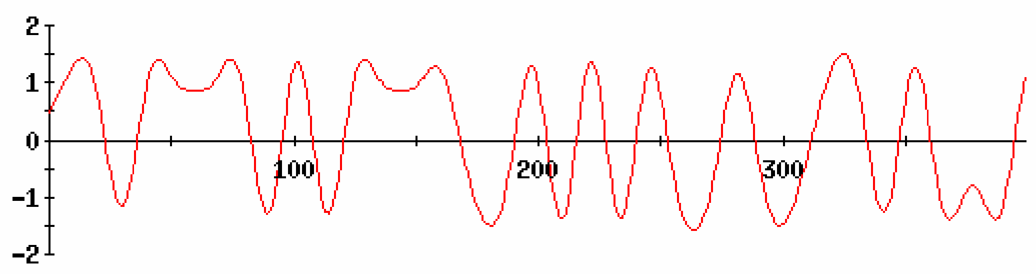


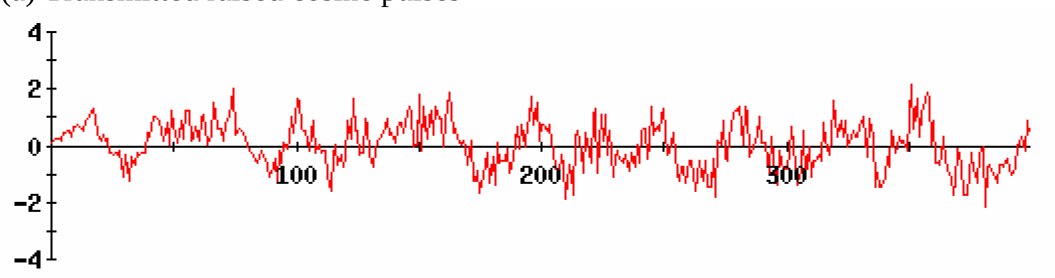
Fig. 13 – A RRC-like symbol and its matched filter output.

I have made up a pulse to approximate the root raised cosine pulse. As before the output of the matched filter is maximum in the middle but fluctuates all over.

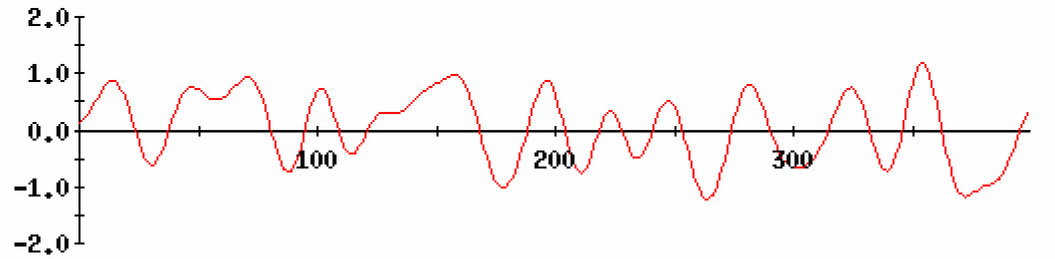
Figure 12 below shows a raised cosine pulse shaped signal. The noise is added to this signal and the received signal is then filtered using a raised cosine filter, which happens to be a matched filter to itself. The resulting pulse is then sampled in the middle of the symbol to decode the bit.



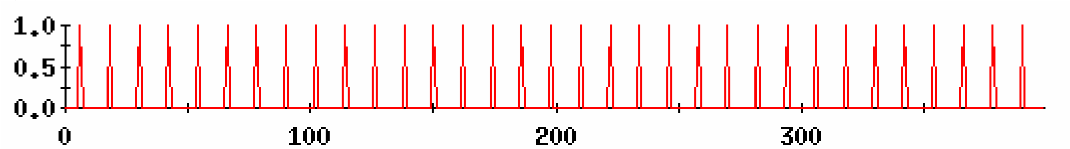
(a) Transmitted raised cosine pulses



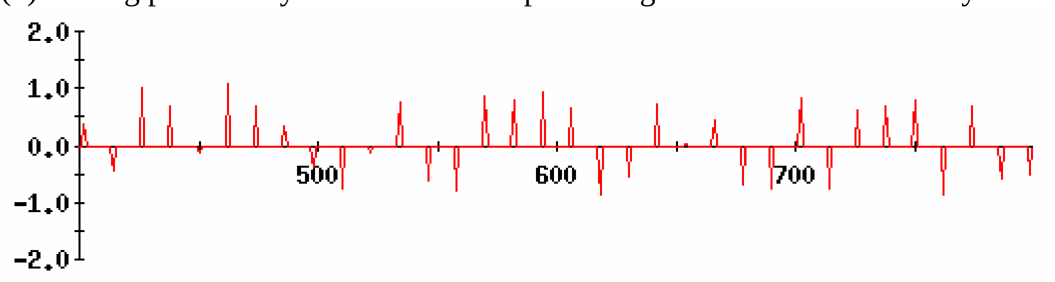
(b) Received signal with noise



c) filtered with a root-raised cosine filter



(d) Timing pulses at symbol rate the sample the signal at the center of the symbol



(e) Center sample, if it is positive decide a 1 bit, if negative decide a 0 bit.
Fig. 14 Root Raised cosine as a matched-pair filter set.

A root raised cosine pair by virtue of its symmetrical shape is a matched filter pair but in addition it also suppresses intersymbol interference so we kill two birds with one stone and that is why this is such a popular filter set.

The optimum filtering again for this type of pulse is the center value as it is for all other non square pulses.

Other non-symmetric pulse

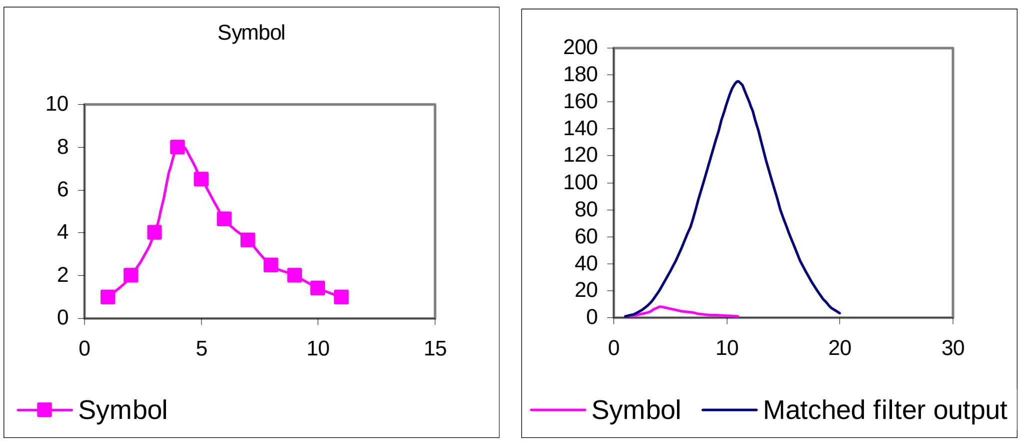


Fig. 15 – An unsymmetrical received signal and its matched filter output.

In the above case, I have a unsymmetrical pulse. Its maximum also occurs at the end of the symbol period, so recognizing that the SNR is not constant throughout the symbol, we would use center sample to decide the fate of the bit.

Summary

1. A matched filter is kind of an averaging filter.
2. It maximizes the signal to noise ratio of a received signal.
3. A matched filter minimizes the BER of the received signal. The BER equations, assume that a matched filter is used to decode the signal.
4. Its time domain shape it the time-reversed replica of the transmitted shape, For symmetrical shapes and most are, the same filter can be used to filter the received signal.
5. Two symmetrical filters make a matched pair, one to transmit, the other to receive.
6. Root raised cosine is a special such pair, because it also reduces inter-symbol interference.
7. Integrate and dump filters is a matched filter only for square pulses, usually used in baseband systems that have no rf carrier.
8. Other channel filtering can make a matched filter non-ideal but is still optimum due to need to build practical system.
9. A matched filter provides us a framework for creating equalized filters (to be covered in a subsequent tutorial.)

Charan Langton
I can be contacted at
mntcastle@earthlink.net

Other tutorials at
www.complextoreal.com