Applied Radio Labs

www.radio-labs.com

Group Delay
Explanations and Applications

Design File: DN004 15 November 1999

Introduction

Group delay is a concept that radio engineers are encountering more frequently as higher performance digital communication systems are developed. It appears that frequency selective components have group delay which, in most cases, seems to degrade system performance. This article addresses several aspects of this topic:

1. Must we have group delay?
2. The relationship between group delay and signal delay.
3. Negative group delay
4. The relationship between group delay and filter loss in passive filters

Definition of Group Delay

Group delay has a mathematical definition (unfortunately), and so we should start with that.

Given a linear system block with frequency domain transfer function $H(j\omega)$

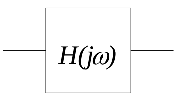

Figure 1 - Linear Component

Writing

$$H(j\omega) = A(j\omega)e^{j\varphi(j\omega)} \qquad (1)$$

then the group delay τ is defined as

$$\tau(\omega) = -\frac{\partial\varphi(\omega)}{\delta\omega} \qquad (2)$$

That is, it is the negative of the rate of change of phase with frequency. The quantity τ has the dimension of time, but the question is; what time does it represent? Many textbooks follow the definition of group delay with a discussion of an ideal element that delays a signal by time T (for example an ideal lossless transmission line). This delay element has a transfer function of

$$H(j\omega) = e^{-j\omega T} \qquad (3)$$

and it is can be seen that the group delay is T for all frequencies. By analogy with this then it is implied that the group delay is in some way related to signal delay. The group delay may also be referred to as "envelope delay".

The Delay of Signals

Given a filter and a plot of group delay, it is natural to assume that this is in some way enables one to determine the delay of signals through the filter. The answer is that group delay and signal delay are related, but not always in an obvious manner.

To start, consider how do we measure delay? The obvious way is to put a pulse into the filter and see how long we have to wait until it comes out. Consider a simple low pass filter:

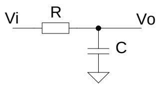

Figure 2 - RC Low Pass Filter

The transfer function is

$$H(j\omega) = \frac{1}{1 + j\omega/\omega_0} \qquad (3)$$

$$\tau(\omega) = \frac{\omega_o}{\omega^2 + \omega_o^2} = \frac{\tau}{1 + \omega^2\tau^2} \qquad (4)$$

with $\omega_o = \frac{1}{RC} = \frac{1}{\tau}$

The group delay at $\omega = 0$ is τ and it gets less at higher frequencies. The group delay is plotted in the following figure for $\tau = 1$.

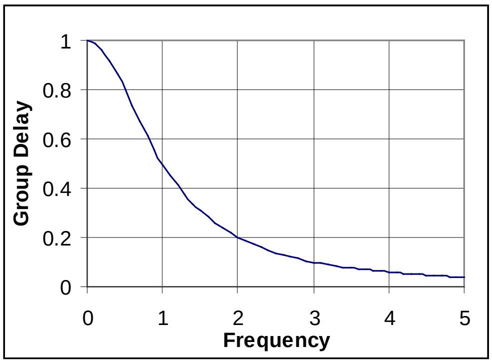

Figure 3 - Group Delay of RC Filter

The rising edge of the input pulse is represented as a step function at *t*=0, and the output waveform is calculated to be

$$v_o(t) = 1 - e^{-t/\tau} = 1 - e^{-\omega_o t} \qquad (5)$$

which is plotted in the following figure, again for the case of $\tau = 1$

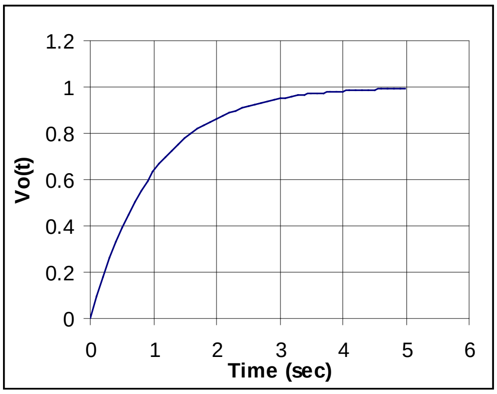

Figure 4 - Step Response of RC Filter

If we define the pulse delay as the delay to the 50% point, this occurs at $t = 0.693\tau$. Can we predict this easily from the group delay graph? Not easily. The signal value at time equal to the DC group delay is 0.632. Maybe we should measure all delays at the 63.2% point?

As in this example, filters often change the shape of signals (this may in fact be the reason it is being used). In general, if the filter significantly changes the shape of the signal, then the delay measurement is somewhat arbitrary and not easily related to the group delay. This does not mean that it is not useful to have specific definitions of delay for particular applications. For example it may be useful to define the delay of a logic filter to be that to the 50% point, as that is the threshold level of a particular logic gate. All we are saying is that this delay is not directly related to the group delay of the filter. Typically a filter used for pulse shaping has specifications on risetime and pulse delay (with specific definition), and these requirements are usually analysed by analysing or simulating the filter in the time domain, not by considering the group delay.

There are more mathematical treatments of pulse transmission, Kuo [2] provides some references.

Thus, group delay is not directly applicable to determining the pulse delay of a low pass filter, unless the pulse shape is passed reasonably undistorted by the filter. The reason for this will become clearer in the next section.

RF Signals and Envelope Delay

RF engineers are more typically concerned with bandpass systems. A typical filter may be used for transmitting one or a number of signals.

For example consider a radio system operating in a 1.25MHz channel at 850MHz (such as IS95). If this signal is passed through a 25MHz bandpass filter in a base station, then it is meaningful to ask what delay the signal experiences. The analogy to the pulse response in the low pass case is the response of the bandpass filter to a burst of carrier. The start of the carrier burst is shown below:

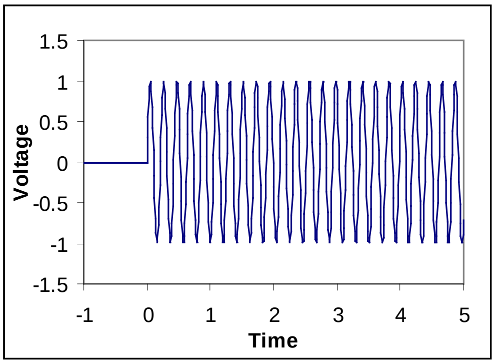

Figure 5 - Start of Carrier Burst

When passed through the filter, the output may be something like

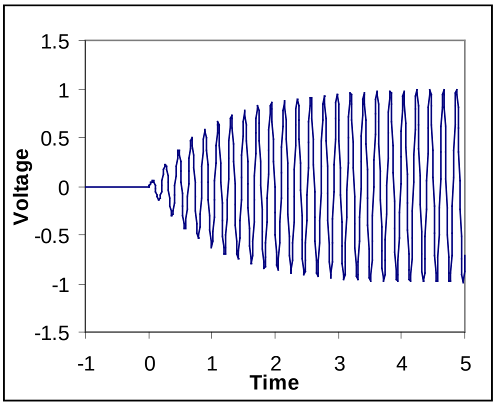

Figure 6 - Filtered RF Burst

The rising edge of the envelope has been delayed, in a similar fashion to the low pass filter. If we analysed this in detail, we would find that the delay depended on the point on the envelop chosen for measurement, and would generally be difficult to relate to the group delay characteristics of the filter. The analogy of the low pass case is that if the envelope undergoes significant change of shape, then the delay is not easily related to the group delay.

When Group Delay equals Envelope Delay

Consider a signal that is bandlimited (i.e. essentially all the power lies within some frequency channel). This signal can be considered a sum of frequency components across this channel. The filter acts on each component separately. If, over this channel, the filter has a constant amplitude response and a constant group delay, then the filter can be replaced by a delay element (and attenuator and a phase shifter).

To summarise, if the filter transfer function is given by (1) and (2) then

if

1. the amplitude response of the filter is approximately constant $\approx A(j\omega_o)$ over the bandwidth of the signal

and

2. the group delay of the filter is approximately constant $\approx \tau(j\omega_o)$ over the bandwidth of the signal

then

3. the output of the filter approximates a replica of the input signal, but time delayed by $\tau(j\omega_o)$, scaled in amplitude by $A(j\omega_o)$ and with a phase shift $\varphi(j\omega_o)$.

A proof of this is given in the appendix.

This is why communications channels have requirements on both the amplitude response and the group delay variation. The amplitude response is often easily associated with the requirement to pass all the signal power. The group delay requirement is to minimise group delay distortion. The total delay is often unimportant (in radio systems it is normally swamped by the propagation delay), however variations in group delay across the channel will cause distortion of the signal waveform.

The reliable determination of group delay requirements for any particular communications system must come from some analysis or simulation or measurement of the system when subjected to group delay variation. These days, simulations are frequently the simplest way of determining the sensitivity to group delay variations.

Must Filters Have Delay

With the exception of the few applications requiring an approximation of a time delay, most engineers would agree that group delay has few redeeming features and it would be nice if we could discover a set of filters that did not exhibit any of this undesirable characteristic. Is this possible?

Clearly we cannot have a filter output appearing before its input, (this would violate causality) so the signal must have a positive (or zero) delay, but why does it have to be large enough to concern system designers?

A filter has to respond differently to different frequencies. What are the difficulties of the job? Try it out – pretend that you are a low-pass filter with the requirement to pass signals below 1000Hz and stop those above 1200Hz. Suppose that you are presented with a signal containing the following sinusoidal components (because you are new to the job I'll split the signal into the sinusoids for you – if you want to do the job on a full-time basis you'll have to learn how to do that yourself!)

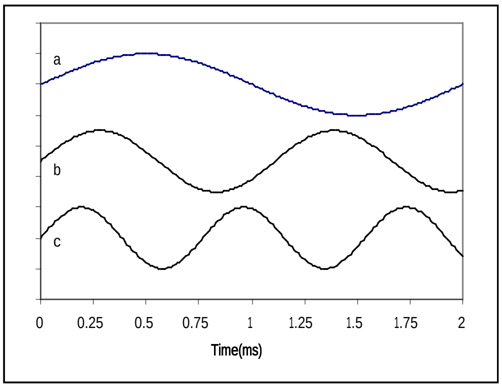

Figure 7 - Input Signals to Filter

Signal 'a' is at 500Hz, signal 'b' at 900Hz and signal 'c' at 1300Hz.

Look at signal 'a'. We are only interested in seeing if signals are below 1kHz to pass, so by 1ms we can be sure that this is a low-frequency signal – we'll pass it.

Signal 'b' is another matter, by 1ms we have almost seen a cycle and we are not 100% that it is below 1kHz, just keep it a bit longer to check, by 2ms we are pretty confident that it is below 1kHz so we can let it out. It is now delayed by 2ms..

Signal 'c' is similar to 'b', except that after we keep it for around 2ms we decide not to let it out.

Now consider a real low pass filter, a 5th order Chebycheff, 1dB ripple, passband edge of 1kHz. The amplitude response is shown below

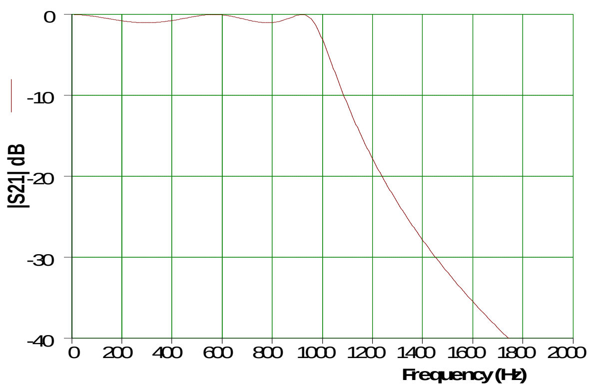

Figure 8 - Amplitude Response 1kHz LPF

The group delay for this filter is shown in the next figure:

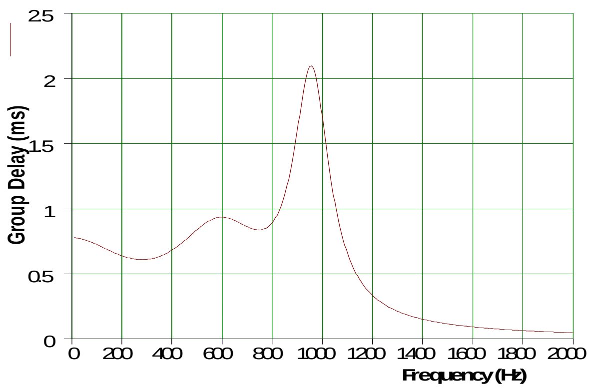

Figure 9 - Group Delay 1kHz LPF

This filter seems to have the same problem as we had, it has to keep signals near the transition region for longer in order to decide whether to pass them or not. It is worth consulting a book of filter tables and seeing how the delay near the passband edge increases as the transition region narrows, i.e. as the filter has to decide between signals closer together in frequency.

Of course there are other, more mathematical and more rigorous ways of showing the relationship between the amplitude response of the filter and the group delay, see for example Papoulis [3].

So we have seen that delay is necessary to allow a filter to actually discriminate between signals of different frequencies. The converse is not true – delay does not imply that the signal will be changed in any way – a piece of ideal transmission line is one counter-example.

The good news is that we have only shown that a filter must have some delay, not that the delay must be variable. Luckily, by equalisation, we can selectively add delay to any filter in order to increase the delay and make it more constant. A well designed equaliser can reduce the delay variations across the filter passband arbitrarily. (Of course there are practical limits to which any filter can be equalised.)

Negative Group Delay

It is possible to come up with circuits that have negative values of group delay. One example is [5]

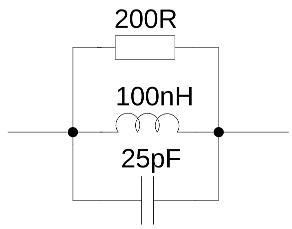

Figure 10 - Notch Filter

This is a notch filter with amplitude and phase responses (in 50Ω system) given below.

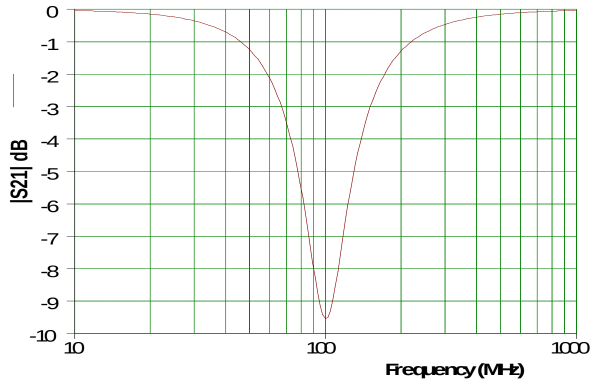

Figure 11 - Amplitude Response of Notch Filter

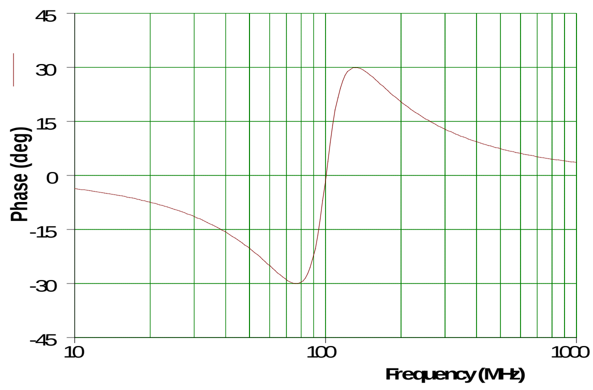

Figure 12 - Phase Response of Notch Filter

Around 100MHz, it can be seen that the phase slope is positive, corresponding to a negative group delay as shown on the group delay plot below:

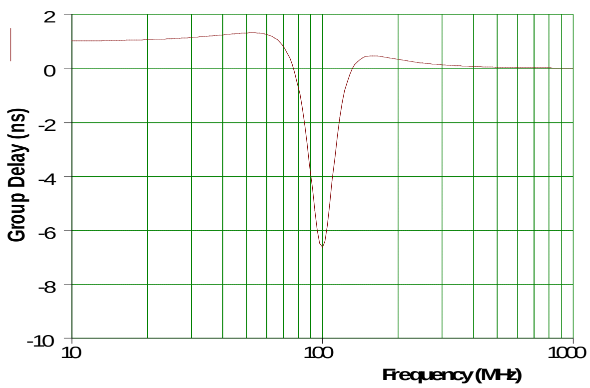

Figure 13 - Group Delay of Notch Filter

At 100MHz the group delay is –6.6ns. Why can't we use this to produce an output from the filter before the input occurs. The reason is that neither the amplitude or group delay are constant in this region. Any pulse put into the filter would suffer significant distortion due to the amplitude and group delay variations, and you can be sure that the total effect is a causal response – output after input.

Group Delay and Filter Loss

We made the intuitive argument before that a filter had to keep a signal for a period of time, in order to decide how to process it. In the case of a digital filter, this storage is easily visualised as sample values in some registers. What about RF filters? They must delay the signal also, and the only way they can achieve this is to actually store the signal energy. This energy is stored in the reactive fields of inductors, capacitors, transmission lines and any other reactive element.

It is in the process of storing this energy that a filter has excess loss. By excess loss we mean the additional attenuation a real filter exhibits when compared to the analysis of an ideal filter with lossless elements.

Bode [4] derived a simple approximation to the excess loss $\Delta L_A(\omega)$ of a real filter (although a more convenient reference for RF engineers is Matthaei, Young & Jones [2] Sec 4.13). A useful form of this suitable for low pass filters and narrow-band bandpass filters is

$$\Delta L_A(\omega) \approx 8.686 \frac{\omega_o}{Q} \tau(\omega) \text{ dB} \qquad (6)$$

here ω_o is the centre frequency of a bandpass filter or the cutoff frequency of a low pass filter, $\tau(\omega)$ is the group delay and Q is the average quality factor of the resonators at ω_o. The derivation of this equation assumes that all reactive elements in the filter have the same Q, however reasonable results are obtained if an average Q factor is used. Hence in an LC filter, reasonable results are obtained using

$$\Delta L_A(\omega) \approx 8.686 \frac{2\omega_o}{Q_L + Q_C} \tau(\omega)$$

Equation (6) is extremely useful in estimating the loss of RF filters. Knowing the Q of resonators, and determining the group delay from filter tables, enables the loss of a filter to be estimated. The results are most accurate if the Q can be estimated from a real filter using (6), then the losses of other filter responses can be estimated.

As an example, consider a 1GHz bandpass filter. We decide that a 4th order Chebycheff is desirable with 0.5dB ripple and 100MHz bandwidth. How high a Q is required for the resonators to keep the excess loss at midband below 0.5dB.

The group delay of the bandpass filter is calculated to be 13.3ns. From (6) we see that the Q required is 1452. The following figure shows the insertion loss of the ideal filter and one using resonators with a Q of 1500. Notice the typical rounding of the filter edges in the lossy case.

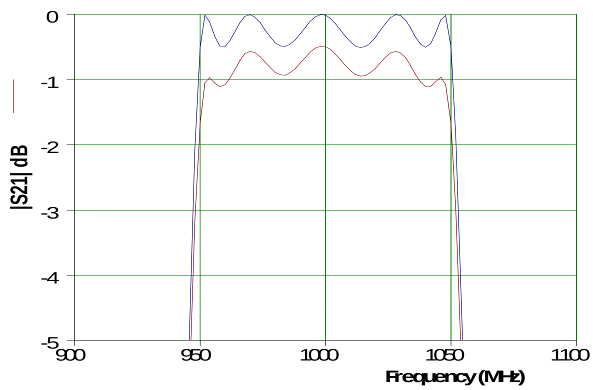

Figure 14 - Insertion Loss of 1GHz BPF

The filter group delay is shown in the next figure:

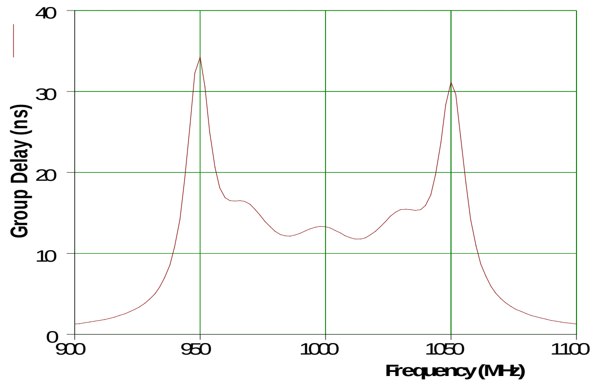

Figure 15 - Group Delay of 1GHz BPF

When the excess loss is determined from Figure 14, and compared with that predicted by equation (6), the results are very similar as shown in the next figure. The only significant approximation in the simulation is that the resonator Q is constant with frequency.

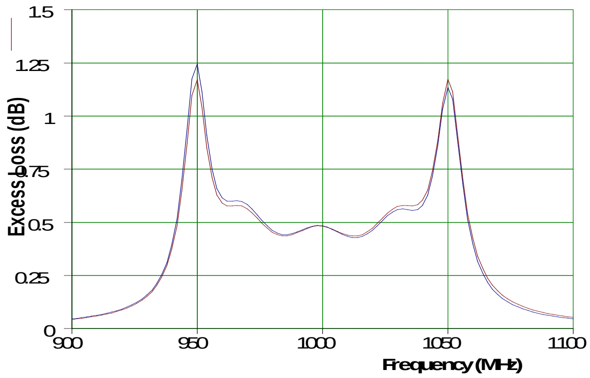

Figure 16 - Excess Loss of BPF, Trace 1 is simulation, Trace 2 is equation (6) (Q = 1500)

So we now see that the group delay determines the excess loss of passive filters. This is why the effect of loss causes the sharp edges of ideal filters to become rounded. A sharp edge is always associated with high group delay, and this causes high excess loss.

Now consider again the circuit with negative group delay. What happens if the reactive elements in that circuit are lossy? Well, assume that the inductor and capacitor both have a Q of 100. Using (6) we obtain an excess loss of -0.36dB at 100MHz (the centre of the notch) i.e. the filter is not as lossy as otherwise. This is what we would expect, the notch depth decreases due to lossy components. Simulating the circuit shows that the notch depth decreases by 0.35dB.

So circuits with negative group delay can be classified as those circuits which show an increase in response when the reactive components are lossy. Bandstop filters are one such class of circuit.

REFERENCES:

1. Microwave Filters, Impedance Matching Networks and Coupling Structures" G L Matthaei, L Young and E M T Jones McGraw-Hill 1964
2. "Network Analysis and Synthesis" F F Kuo, Wiley 1966
3. "Signal Analysis" A Papoulis McGraw-Hill 1977
4. "Network Analysis and Feedback Amplifier Design" H W Bode Van Nostrand 1945
5. Randy Rhea, RF Globalnet Active Devices Forum, 9th November 1999.

Appendix

Given a bandlimited signal $v_i(t)$ with Fourier Transform $V_i(j\omega)$ which is constrained into the frequency band $\omega_0 \pm \delta\omega$, then when this signal is passed through the filter in Figure 1.

$$V_o(j\omega) = V_i(j\omega)H(j\omega)$$

If the amplitude response of the filter is approximately constant over the bandwidth of the signal i.e.

$$A(j\omega) \approx A(j\omega_o) \text{ for } |\omega - \omega_o| < \Delta$$

and the group delay of the filter is approximately constant over the bandwidth of the signal i.e.

$$\tau(j\omega) \approx \tau(j\omega_o) \text{ for } |\omega - \omega_o| < \Delta$$

then the transfer function can be written

$$H(j\omega) \approx A(j\omega_o)e^{-j(\varphi(\omega_o) - \tau(\omega_o)(\omega - \omega_o))} \text{ for } |\omega - \omega_o| < \Delta$$

where a first order Taylor expansion has been performed on the phase. The fourier transform of the output of the filter can be written

$$V_o(j\omega) = V_i(j\omega)\left[e^{-j\tau(\omega_o)}\right]\left[A(j\omega_o)\right]\left[e^{-j(\varphi(\omega_o) + \tau(\omega_o)\omega_o)}\right] \quad \text{(A1)}$$

valid for all ω as $V_o(j\omega)$ vanishes outside $|\omega - \omega_o| < \Delta$. The first bracketed term in (A1) is seen to be a delay $\tau(j\omega_o)$, the second an amplitude scale of $A(j\omega_o)$ and the third a phase shift ensuring that the total phase shift at ω_o is $\varphi(j\omega_o)$.

Peter White
Applied Radio Labs

For further information visit www.radio-labs.com
Comments and criticisms are welcome.
Email: feedback@radio-labs.com