

Calculating the transfer functions of the filters/crash course in network theory

A.1. Modelling the circuit

Calculating the transfer of an electric or electronic filter usually goes in two steps:

1. Draw a network model, that is a mathematical abstraction of the real-life circuit
2. Calculate the transfer of this model

Regarding the first step: the simplest model for a capacitor is an ideal linear capacitor (which my former network theory professor Fred Neerhoff always called a *lineaire capaciteit*, a linear capacity), the simplest model for a resistor is an ideal linear resistor, the simplest model for an inductor is an ideal linear inductor (linear inductivity), the simplest model for an amplifier with well-defined gain is a linear controlled source, the simplest model for a high-gain amplifier used in a feedback configuration is a nullor (combination of a nullator and a norator, see below) and the simplest model for an interconnecting wire is a node.

If you are worried about second-order effects, you need to use more elaborate models. To give two examples that are relevant to audio: an inductor in a passive crossover could be represented by the series connection of a linear inductivity and an ideal linear resistor if you are worried about the effect of its DC resistance. An amplifier could be represented by an ideal controlled source plus a resistor representing its output resistance, if there is reason to believe that the output resistance might not be negligible. In this appendix, we will only use ideal models with no second-order effects.

A.2. Transfer functions

The transfer of a filter can be described by a transfer function consisting of the ratio between two polynomials in s , which in older literature is denoted as p . (This actually only applies when the filter is linear, time invariant, continuous time and lumped, but analogue filters used for audio are usually close enough to being all of that to model them with a linear, time-invariant, continuous-time lumped network). Personally I think the old notation p is much clearer than s , because s is too similar to s , the SI symbol for second. Still, since s is the more usual notation, I will stick to it.

Depending on the type of calculation one wants to do, s can be regarded as the Laplace variable (outside the scope of this article), as a differentiation to time operator, or as $j\omega$, where ω is the radian frequency ($\omega = 2\pi f$) and j is the imaginary unit number ($j^2 = -1$).

For example, suppose the transfer function of the filter is of the form:

$$H(s) = \frac{V_{\text{out}}}{V_{\text{in}}} = K \frac{a_2 s^2 + a_1 s + 1}{b_2 s^2 + b_1 s + 1}$$

Interpreting s as a differentiation-to-time operator, this means that the relation between the input and output voltage is given by this differential equation:

$$b_2 \frac{d^2 v_{out}}{dt^2} + b_1 \frac{dv_{out}}{dt} + v_{out} = K \left(a_2 \frac{d^2 v_{in}}{dt^2} + a_1 \frac{dv_{in}}{dt} + v_{in} \right)$$

Using complex numbers, calculating the output signal becomes relatively simple when the filter does not oscillate and when the input signal is a stationary sine or cosine wave (that is, a sine or cosine that has been there long enough for initial transients to damp out). A cosine equals the sum of two complex exponential functions:

$$\cos(\omega t) = \frac{1}{2} e^{j\omega t} + \frac{1}{2} e^{-j\omega t}$$

As the network is linear, we may use the superposition principle. That is, we may calculate the response to each complex exponential signal independently and simply add the results.

The time derivative of a complex exponential signal is:

$$\frac{de^{j\omega t}}{dt} = j\omega e^{j\omega t}$$

Hence, when

$$v_{in} = \frac{1}{2} e^{j\omega t}$$

we get

$$b_2 \frac{d^2 v_{out}}{dt^2} + b_1 \frac{dv_{out}}{dt} + v_{out} = K (a_2 (j\omega)^2 + a_1 j\omega + 1) \frac{1}{2} e^{j\omega t}$$

This equation is satisfied when the output signal is also a complex exponential signal of the same frequency, but multiplied with some complex multiplication factor. That is, assume that

$$v_{out} = X e^{j\omega t}$$

This results in

$$(b_2 (j\omega)^2 + b_1 j\omega + 1) X e^{j\omega t} = K (a_2 (j\omega)^2 + a_1 j\omega + 1) \frac{1}{2} e^{j\omega t}$$

$$X = \frac{1}{2} \cdot \frac{K (a_2 (j\omega)^2 + a_1 j\omega + 1)}{b_2 (j\omega)^2 + b_1 j\omega + 1} = \frac{1}{2} H(j\omega)$$

and

$$\frac{v_{out}}{v_{in}} = \frac{X}{\frac{1}{2}} = H(j\omega)$$

So when you substitute $s=j\omega$, the transfer function turns into a complex-valued gain factor for complex exponential input signals. Note that $H(j\omega)$ can also be written in a polar form:

$$H(j\omega) = |H(j\omega)|e^{j\phi}$$

with

$$\phi = \arctan(\operatorname{Im}(H(j\omega)) / \operatorname{Re}(H(j\omega))) + k\pi$$

when the real part of $H(j\omega)$ is not zero and where k is an integer. The factor $|H(j\omega)|$ represents the actual gain, while the factor $e^{j\phi}$ just gives a phase shift of ϕ .

In the end, we are interested in the response to the real-valued cosine wave rather than to the complex exponential waveform. Hence, when

$$v_{\text{in}} = \cos(\omega t) = \frac{1}{2}e^{j\omega t} + \frac{1}{2}e^{-j\omega t}$$

then

$$v_{\text{out}} = \frac{1}{2}e^{j\omega t} H(j\omega) + \frac{1}{2}e^{-j\omega t} H(-j\omega)$$

It can be shown that $H(j\omega)$ and $H(-j\omega)$ must have equal real parts and opposite imaginary parts for any filter that produces a real-valued output signal for each real-valued input signal. This results in equal magnitudes, but opposite phases for $H(j\omega)$ and $H(-j\omega)$. Hence,

$$v_{\text{out}} = \frac{1}{2}e^{j\omega t} |H(j\omega)|e^{j\phi} + \frac{1}{2}e^{-j\omega t} |H(j\omega)|e^{-j\phi} = |H(j\omega)| \frac{1}{2}(e^{j(\omega t + \phi)} + e^{-j(\omega t + \phi)}) = |H(j\omega)| \cos(\omega t + \phi)$$

That is, also for a normal cosine wave $|H(j\omega)|$ represents the gain, while the phase shift is ϕ .

A.3. Poles and zeros

The values of s for which the denominator of the transfer function is zero are called the poles of the transfer function. The values of s for which the numerator of the transfer function is zero are called the zeros. The number of poles of a lumped system can never exceed the number of energy-storing parts (such as capacities and inductivities in an electric network, or masses and springs in a mechanic system). The denominator of the transfer function is also known as the characteristic equation.

An interesting property is that all transfers of a system with multiple in- or outputs have the same poles, although in some of these transfers some poles may be covered by zeros.

A.4. Impedance of ideal capacitors and inductors

The current through a linear capacity is C times the derivative of the voltage across it to time. Hence,

$$I = s C V$$

so the impedance is

$$Z = \frac{V}{I} = \frac{1}{sC}$$

The voltage across a linear inductivity is L times the derivative of the current through it to time, so its impedance is

$$Z = sL$$

A.5. Calculating the transfer function for a given filter network

The transfer of any network consisting of impedances, independent sources, linear controlled sources, nullators and norators can always be calculated with a method called modified nodal analysis. The procedure is as follows:

A. For a network with n nodes, number the nodes from 0 up to and including $n - 1$. Node 0 will be the reference node (also known as the datum), all voltages are with respect to node 0. The ground node is usually taken as node 0, although this is not necessary (in fact the term ground has no meaning in network theory).

B. For each of the $n - 1$ nodes that have a number different from 0, write down the nodal equation. This is an equation expressing how much current flows into and out of the node as a function of the node voltages. Of course all current that flows into a node also has to flow out of the node again (Kirchhoff's current law).

C. For each voltage source, each controlled source and each nullator or norator, you have to introduce additional equations.

D. The last step is to solve the unknown node voltages from the resulting set of equations. With linear equations there is a simple trick for this, which is known as Gaussian elimination. The trick is best illustrated by an example. Suppose you have the following set of equations:

$$\begin{aligned} 3a + 2b + 3c &= 0 \\ 2a + 5b + 8c &= 0 \\ a + 6b + 4c &= 12 \end{aligned}$$

By subtracting two thirds of the first equation from the second equation and one third of the first equation from the third equation, you can eliminate a from the second and third equations:

$$\begin{aligned} 3a + 2b + 3c &= 0 \\ \frac{11}{3}b + 6c &= 0 \\ \frac{16}{3}b + 3c &= 12 \end{aligned}$$

By subtracting 16/11 times the second equation from the third equation, b is eliminated from the third equation and we get:

$$\begin{aligned} 3a + 2b + 3c &= 0 \\ \frac{11}{3}b + 6c &= 0 \end{aligned}$$

$$-\frac{63}{11}c = 12$$

The third equation now has only one unknown left (c) and is quite easily solved. Once the third equation is solved, you can substitute the result in the second equation, which then also becomes a simple equation with only one unknown (b). Substituting b and c in the first equation then results in a simple equation with only the unknown a .

$$\begin{aligned} -\frac{63}{11}c = 12 &\Leftrightarrow c = -\frac{11 \cdot 12}{63} = -\frac{132}{63} = -\frac{44}{21} \\ \frac{11}{3}b + 6 \cdot \left(-\frac{44}{21}\right) &= 0 \Leftrightarrow b = \frac{3}{11} \cdot 6 \cdot \frac{44}{21} = \frac{3 \cdot 6 \cdot 4}{21} = \frac{24}{7} \\ 3a + 2 \cdot \frac{24}{7} + 3 \cdot \left(-\frac{44}{21}\right) &= 0 \Leftrightarrow a = \frac{1}{3} \left(-2 \cdot \frac{24}{7} + 3 \cdot \frac{44}{21}\right) = \frac{1}{3} \left(-\frac{4}{7}\right) = -\frac{4}{21} \end{aligned}$$

It is clear that the amount of work increases rapidly with increasing number of equations (though not as rapidly as with other algorithms for solving systems of linear equations). As the number of equations in modified nodal analysis depends on the number of nodes, it is advisable to eliminate any node that can easily be eliminated before doing the analysis.

Modified nodal analysis is a method that is typically only used by humans (as opposed to computers) when they don't see a simpler way to calculate the transfer. For example, when two impedances form a voltage divider, you can just use the equation for a voltage divider instead of going for a full-blown modified nodal analysis.