Analysis of the Moog Transistor Ladder and Derivative Filters

Timothy E. Stinchcombe [†]

25 Oct 2008

[†]Comments, suggestions and corrections can be emailed to me at:
tim102@tstinchcombe.freeserve.co.uk.
A good place to seek answers to questions on the internals of synthesizers in general is the 'Synth DIY' mailing list: http://dropmix.xs4all.nl/synth-diy/

Contents

1 Introduction **4**

2 Moog Transistor Ladder Filter **5**

2.1 Large-signal development of the transfer function 5

2.1.1 The differential pair 5

2.1.2 The filter core 8

2.2 Small-signal development of the transfer function 13

2.3 Adding a feedback loop 16

2.4 Accuracy of the transistor ladder model 17

2.5 Poles and frequency responses 19

3 Diode Ladder Filters **25**

3.1 Introduction 25

3.2 Transfer function derivations 28

3.3 Adding the feedback loops 35

3.4 Accuracy of the diode ladder models 36

3.5 Poles and frequency responses 39

3.6 Later EMS 5-pole filter with zero 48

4 Conclusions **50**

1 Introduction

After the introduction of its design in the mid-60s, the Moog transistor ladder filter [1, 2] quickly became a favourite, and is something of a benchmark against which other filters used in synthesizers tend to be judged. Much of the usefulness of the filter is the ability to apply *voltage-control* to the cut-off frequency, as it was this, along with voltage-controlled oscillators and amplifiers, that helped turn the electronic music synthesizer into a much more practical musical instrument than it had hitherto been. Its main endearing feature is, of course, that it produces a wide range of pleasant sounds!

The Moog ladder filter has undoubtedly been the subject of a great deal of analysis over the intervening decades, but very little of this seems to have appeared in print. It was this paucity of available analyses of the filter that initially led me to working out what I wanted to know for myself, resulting in [3]: the only other paper I could find at the time was [4], but this gives little insight as to how the physical structure of the filter actually results in any filtering action. Since that time though (early 2004), several papers relating to the Moog ladder filter and its derivatives *have* appeared, [5, 6, 7, 8], but which all concentrate on *digital* implementations of the filter.

Moog patented the filter, specifying *transistors* as the active element in the ladder, and so it was probably inevitable that sooner or later derivative filters using *diodes* in the ladder would appear, which would thus not infringe the patent. The basic principles of how the filter works remains roughly the same using diodes, but there are differences, and exploitation of these appears to have led to a larger variety of diode ladder filters than the original Moog design, which has tended to remain fairly static. The differences extend into how the filters sound of course, and discerning ears are no doubt capable of distinguishing the filter types from each other.

The main aim of this paper is thus to put on record some of the ways both types of filter can be analysed, to look at some of the commonalities and differences, and to examine the usual characteristics associated with such filters, for example the poles and magnitude and phase responses etc. The main emphasis is on doing this *just for its own sake*, rather than trying to read too much into the results in terms of how the differences in the analysis might translate into how the filters sound (though this is hard to resist on occasion!). Whilst most here is the original work of the author, there isn't anything which is that remarkable, and it is intended to be a compendium of information above anything else—at the very least, some of the equations may give some assistance in choosing capacitor values given the desired frequency range and expected ladder currents, or how much gain to provide around the loop to ensure self-oscillation.

The first half of the paper concentrates on the original Moog ladder, initially deriving the transfer function from a large-signal standpoint by looking at the differential

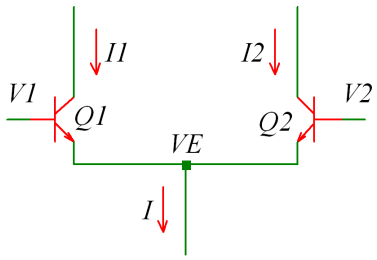


Figure 1: The differential, or long-tailed, pair

pair in some detail. The 'large-signal' moniker presents something of a quandary, as inevitably one makes approximations which move the analysis into the small-signal arena. However it does seem useful as a means to distinguish it from the quite different route using the simplified hybrid-π small-signal model which follows it. After examining the characteristics of this filter, the same pattern is followed for the diode ladder filters (with the exception that there is no 'small-signal' treatment).

2 Moog Transistor Ladder Filter

2.1 Large-signal development of the transfer function

In this section we derive the transfer function of the basic 'core' of the filter, by first analysing the behaviour of the standard differential pair. The overall transfer function of the filter *including* a feedback path is derived later in Section 2.3.

2.1.1 The differential pair

The relationship between the collector currents and differential voltage input to a differential or long-tailed pair plays a key role in the following sections which derive the various filter transfer functions. Thus in this section we take some time to derive and examine this relationship: I find the analysis that utilizes the hyperbolic tangent to be particularly succinct and elegant, as can be found in [9], but much of the missing detail has been added here.

Suppose we have a pair of NPN transistors tied at the emitter, as shown in Figure 1: the common emitter voltage is V_E, and from which point we draw current I. The base voltages are V_1 and V_2, and the collector currents are I_1 and I_2, which we also assume are the emitter currents (i.e. we neglect the base currents). Using a standard simplification of the Ebers-Moll model (see for example [10]), with I_s the saturation current as normal, and writing kT/q as V_T for convenience, we have

$$I_1 \approx I_s e^{\frac{V_1 - V_E}{V_T}} \quad \text{and} \quad I_2 \approx I_s e^{\frac{V_2 - V_E}{V_T}}, \tag{1}$$

which upon dividing gives

$$\frac{I_2}{I_1} = \frac{e^{\frac{V_2 - V_E}{V_T}}}{e^{\frac{V_1 - V_E}{V_T}}} = e^{\frac{V_2 - V_1}{V_T}}, \tag{2}$$

and of particular note here is the fact that throughout this paper we will be assuming that our transistors and diodes are perfectly matched, and so their saturation currents, I_s, are identical. Substituting for I_2 from the above in

$$I = I_1 + I_2$$

gives

$$I = I_1 + I_1 e^{\frac{V_2 - V_1}{V_T}}, \tag{3}$$

resulting in

$$I_1 = \frac{I}{1 + e^{\frac{V_2 - V_1}{V_T}}}. \tag{4}$$

Now we require the use of a subtle bit of hindsight, using the sort of trick that students always decry—multiply top and bottom by 2, and 'add zero' into the numerator:

$$I_1 = \frac{I \times 2}{2\left(1 + e^{\frac{V_2 - V_1}{V_T}}\right)} = \frac{I}{2}\left[\frac{1 + 1 + e^{\frac{V_2 - V_1}{V_T}} - e^{\frac{V_2 - V_1}{V_T}}}{1 + e^{\frac{V_2 - V_1}{V_T}}}\right] = \frac{I}{2}\left[1 + \frac{1 - e^{\frac{V_2 - V_1}{V_T}}}{1 + e^{\frac{V_2 - V_1}{V_T}}}\right].$$

Now

$$\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{e^{2x} - 1}{e^{2x} + 1},$$

from which it can be seen we get

$$I_1 = \frac{I}{2}\left[1 - \tanh\left(\frac{V_2 - V_1}{2V_T}\right)\right] = \frac{I}{2}\left[1 + \tanh\left(\frac{V_1 - V_2}{2V_T}\right)\right]. \tag{5}$$

For I_2 we have

$$I_2 = I - I_1 = I - \frac{I}{2}\left[1 + \tanh\left(\frac{V_1 - V_2}{2V_T}\right)\right] = \frac{I}{2}\left[1 - \tanh\left(\frac{V_1 - V_2}{2V_T}\right)\right],$$

and the symmetry between the currents is very obvious: current I splits equally between the two transistors, but if the base voltage at either is higher than the other, than that transistor switches on more, taking incrementally more than its half-share (given by the *tanh* term), whilst the current in the other is decremented by the same amount.

The Taylor series expansion of $\tanh x$ is

$$\tanh x = x - \frac{x^3}{3} + \frac{2x^5}{15} - \frac{17x^7}{315}\cdots \qquad |x| < \frac{\pi}{2},$$

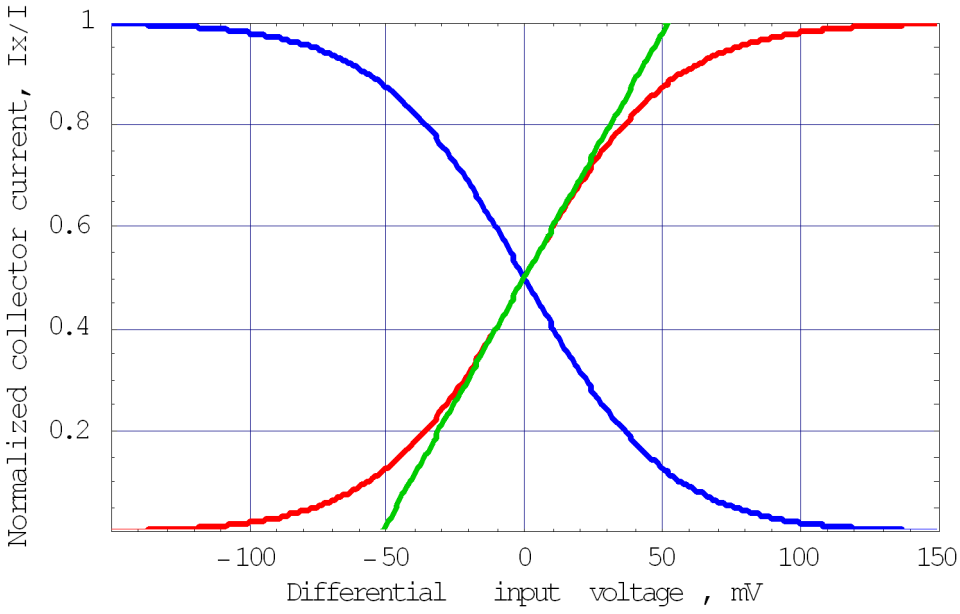


Figure 2: Normalized collector currents, and a linear approximation

so when the differential input voltage is small, say within ±20mV[†], we can make a linear approximation

$$\tanh x \approx x,$$

so the collector currents become:

$$I_1 = \frac{I}{2}\left[1 + \frac{V_1 - V_2}{2V_T}\right], \tag{6}$$
$$\text{and} \quad I_2 = \frac{I}{2}\left[1 - \frac{V_1 - V_2}{2V_T}\right].$$

In the following sections we shall often be interested in the *difference* between these currents, and so simple subtraction yields

$$\begin{aligned} I_1 - I_2 &= \frac{I(V_1 - V_2)}{2V_T} \\ &= 19.2\, I(V_1 - V_2), \end{aligned} \tag{7}$$

where the constant $1/(2V_T)$ evaluates to 19.2 with $V_T = 26\text{mV}$. This last expression effectively gives the transconductance of the differential pair, i.e. i_{out}/v_{in} in terms of the biasing current I, and something like it is frequently seen on operational transconductance amplifier (OTA) datasheets, where I is normally 'I_{abc}', the *amplifier bias current.*

If we plot the collector currents (equations (5)), we get the familiar 'bow tie' pattern—Figure 2 shows this, but to remove the absolute value of I the curves have

[†]We're not going to be too rigorous here, suffice it to say that we just make the approximation 'good enough for most practical purposes'!

been normalized by plotting I_1/I (red) and I_2/I (blue); also shown is the linear approximation to the current on one side, equation (6), to give a visual indication of how close this approximation actually is to the original curve when the differential input voltage is small (the green trace, which is again appropriately normalized).

In reviewing the above working we notice that few conditions are needed to make it usefully general: if the ratio of two entities, generally currents, is the exponential of a constant multiple of a third, generally a voltage difference,

$$\frac{I_2}{I_1} = e^{\frac{-\Delta V}{V_T}},$$

and if the currents sum to

$$I = I_1 + I_2,$$

and $\Delta V/V_T$ is small, then the *difference* in the currents, $\Delta I = I_1 - I_2$ is given by

$$\Delta I = \frac{I\Delta V}{2V_T}. \tag{8}$$

In particular, if the emitters of the transistors in the pair are not tied together so that their emitter voltages are not equal, i.e. we have $V_{E_1} \neq V_{E_2}$, and if the collector currents still sum to give I, then equation (2) becomes

$$\frac{I_2}{I_1} = \frac{e^{\frac{V_2 - V_{E_2}}{V_T}}}{e^{\frac{V_1 - V_{E_1}}{V_T}}} = e^{\frac{V_2 - V_1 - (V_{E_2} - V_{E_1})}{V_T}},$$

and so we immediately get

$$\Delta I = \frac{I(\Delta V - \Delta V_E)}{2V_T} \tag{9}$$

from (8), where $\Delta V = V_1 - V_2$ is the difference in base voltages, $\Delta V_E = V_{E_1} - V_{E_2}$ is the difference in emitter voltages, and $\Delta I = I_1 - I_2$ the difference in collector currents. We will make repeated use of this expression.

2.1.2 The filter core

The basic filter set-up shown in Figure 3, as per Robert Moog's original patent [2], consists of: 'driver transistors' $Q1$ and $Q2$, to which the differential input voltage is applied, and from which current I_f, proportional to the cut-off frequency, is drawn; a pair of 'output coupling transistors', $Q11$ and $Q12$, from which the differential output voltage is taken; and in between, four filter stages, each consisting of a pair of transistors with a capacitor tied between their emitters. The resistor chain biases the transistors so that they are well separated and cannot become saturated, and as such play no role in the filtering action. In this section we are concerned only with the 'core' of the filter:

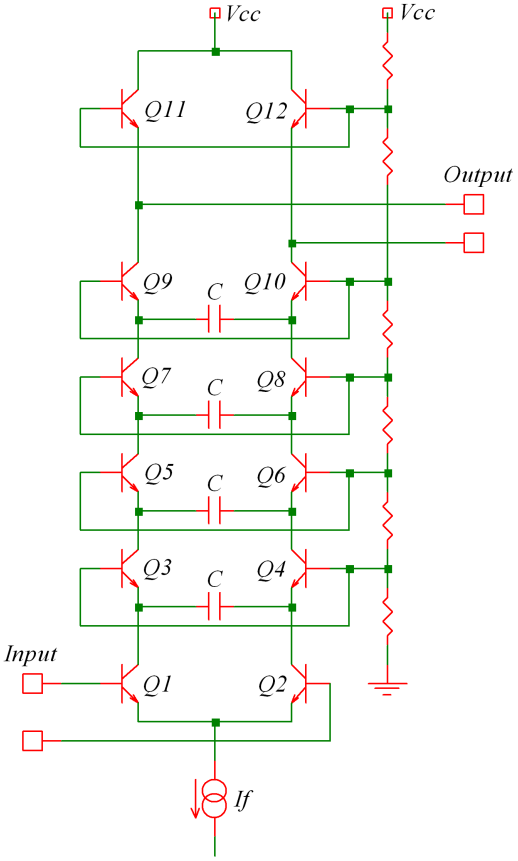


Figure 3: Basic circuit of the Moog ladder filter

later in Section 2.3 we will add a feedback loop and amend the overall transfer function accordingly.

Throughout we assume that the base currents are negligible, and thus that through each transistor the emitter current equals the collector current, and also that the transistors are perfectly matched so that the approximations made in the previous section hold.

With so many currents and voltages involved for each of the stages, the notation for analysing a circuit such as this poses something of a problem: since we will be mainly interested in *differences* in voltages and currents between the sides of the ladder, I have tried to chose a notation which reduces the number of symbols used, and so we move to it as quickly as possible. Figure 4 shows the notation used about stage n: currents in the ladder arms *below* the stage capacitor C are I_{nl} and I_{nr} ('l' and 'r' for left and right), and so their difference is

$$\Delta I_n = I_{nl} - I_{nr}.$$

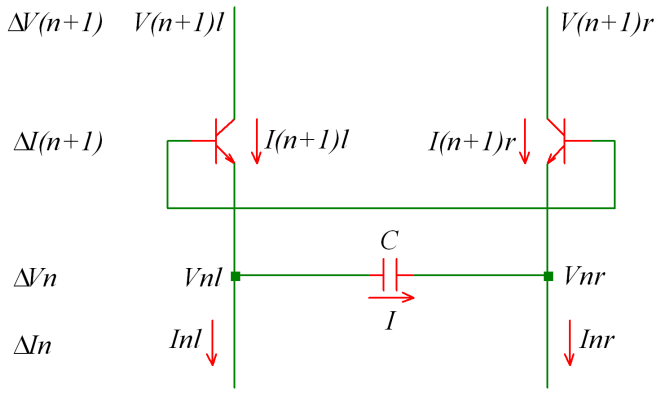


Figure 4: Notation for transistor ladder stage n

The voltages either side of the capacitor are V_{nl} and V_{nr}, with difference

$$\Delta V_n = V_{nl} - V_{nr},$$

and the currents *above* the capacitor, which pass through their respective transistors (recall we are assuming the collector current equals the emitter current), are $I_{(n+1)l}$ and $I_{(n+1)r}$, so

$$\Delta I_{n+1} = I_{(n+1)l} - I_{(n+1)r}.$$

Figure 5 shows the numbering of the stages against the basic core of the filter we are going to use: current I_f, which sets the cut-off frequency, is drawn out of the bottom of the ladder, and since no current can enter or leave the ladder except at the bottom or top (as we assuming that the base currents are negligible), we must have

$$I_{nl} + I_{nr} = I_f$$

for all n. I have also made another slight simplification by removing the top-most pair of transistors, and have instead taken the output from the last filter section—it is easy to see that as the currents through such a top pair will be the same as through the last filter section, the differential voltage at their emitters must also be the same as that across the last filter section, and thus (analytically at least), removing them doesn't change anything (but whether doing so has implications of impedance/loading of the last filter section by the take-off circuitry is a different matter!).

To derive the transfer function of a single stage, we have from Figure 4 that

$$I_{(n+1)l} = I + I_{nl},$$
$$I_{(n+1)r} + I = I_{nr},$$

which subtract to give

$$I_{(n+1)l} - I_{(n+1)r} = \Delta I_{n+1} = I + I_{nl} + I - I_{nr} = \Delta I_n + 2I,$$

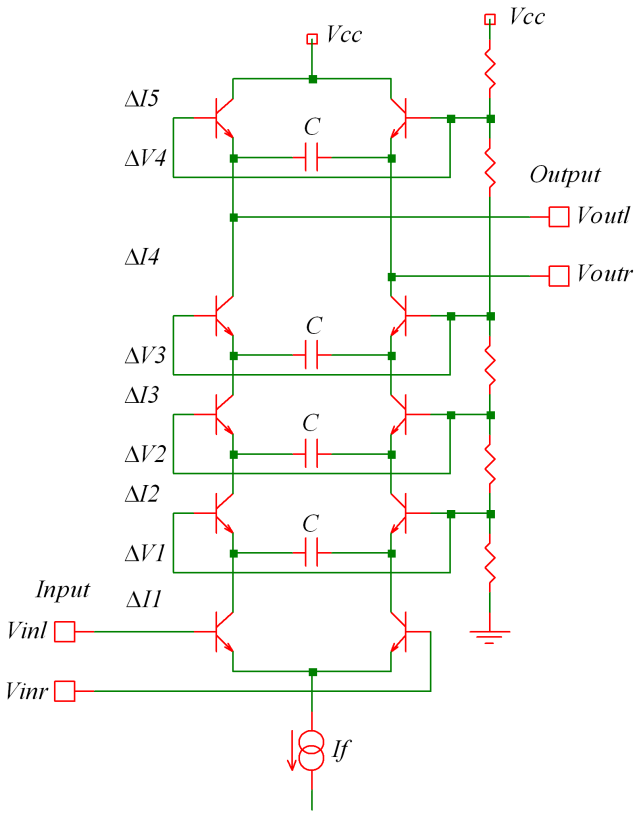


Figure 5: Basic Moog ladder filter core and stage numbering

that is

$$\Delta I_{n+1} = \Delta I_n + 2I.$$

At the capacitor we also have

$$V_{nl} - V_{nr} = \Delta V_n = \frac{I}{sC},$$

so substituting for I in the above gives

$$\Delta I_{n+1} = \Delta I_n + 2sC\Delta V_n. \tag{10}$$

Since $I_{(n+1)l} + I_{(n+1)r} = I_f$, the conditions for using (9) are met, noting that the base voltages are equal, and the difference in emitter voltages is $\Delta V_n = V_{nl} - V_{nr}$, so we have

$$\Delta I_{n+1} = \frac{-I_f \Delta V_n}{2V_T}.$$

Now we can rid ΔV_n from the above, giving

$$\Delta I_{n+1} = \Delta I_n - 2sC\frac{2V_T\Delta I_{n+1}}{I_f},$$

which simply re-arranges to give

$$\frac{\Delta I_{n+1}}{\Delta I_n} = \frac{1}{s\frac{4CV_T}{I_f} + 1}. \tag{11}$$

This is the transfer function of a single stage. If we put $R_{equiv} = 4V_T/I_f$, we could write this as

$$\frac{\Delta I_{n+1}}{\Delta I_n} = \frac{1}{sCR_{equiv} + 1},$$

or better still as

$$\frac{\Delta I_{n+1}}{\Delta I_n} = \frac{1}{\frac{s}{\omega_c} + 1}, \tag{12}$$

where

$$\omega_c = \frac{1}{CR_{equiv}} = \frac{I_f}{4CV_T},$$

or

$$f_c = \frac{1}{2\pi CR_{equiv}} = \frac{I_f}{8\pi CV_T}. \tag{13}$$

This is good for $n = 1, 2, 3, 4$, so to complete the transfer function, we need to look at the input and output stages.

For the input driver pair let the input voltage be $V_{in} = V_{inl} - V_{inr}$, and with the collector currents summing as $I_{1l} + I_{1r} = I_f$, we can apply (7) directly to get

$$\Delta I_1 = \frac{I_f V_{in}}{2V_T}. \tag{14}$$

At the output at the top of the ladder let

$$V_{out} = V_{outl} - V_{outr} \equiv V_{4l} - V_{4r} = \Delta V_4.$$

Similar to stage n above, we have $I_{5l} + I_{5r} = I_f$, and can use (9) in the same manner to immediately get

$$\Delta I_5 = \frac{-I_f \Delta V_4}{2V_T} = \frac{-I_f V_{out}}{2V_T}. \tag{15}$$

For the complete transfer function, divide the V_{out} and V_{in} expressions just derived, and insert $1 = \Delta_n/\Delta_n$ for $n = 2, 3, 4$ repeatedly, and then substitute for the single-stage functions from equation (12):

$$\frac{\frac{-I_f V_{out}}{2V_T}}{\frac{I_f V_{in}}{2V_T}} = \frac{\Delta I_5}{\Delta I_1}, \tag{16}$$

which when cancelled gives

$$\begin{aligned}\frac{V_{out}}{V_{in}} &= \frac{-\Delta I_5}{\Delta I_1} \\ &= \frac{-\Delta I_2}{\Delta I_1} \times \frac{\Delta I_3}{\Delta I_2} \times \frac{\Delta I_4}{\Delta I_3} \times \frac{\Delta I_5}{\Delta I_4} \\ &= \frac{-1}{\left(\frac{s}{\omega_c}+1\right)^4}.\end{aligned}$$

As is usual practice, we normalize this by substituting $s' = s/\omega_c$, to get

$$\frac{V_{out}}{V_{in}} = \frac{-1}{(s'+1)^4},$$

and then simply drop the prime to finally get the normalized transfer function of the filter core as

$$G(s) = \frac{-1}{(s+1)^4}. \tag{17}$$

We will look at some of the characteristics of this transfer function in section 2.5, after having added a feedback loop around it in Section 2.3. However, first we look at a slightly different way of achieving the same end result.

2.2 Small-signal development of the transfer function

In this section we shall develop the transfer function by a slightly different route, starting from small-signal representations of the filter stages. I find this process to be less intuitive and less edifying than the previous method, as the act of setting DC sources etc. to zero seems to make the analysis more removed from reality in my view. It does have the advantage that the approximations we had to work quite hard for above arrive 'pre-packaged' in the model we use—this can also be seen as a disadvantage, as you have little visibility of just *what* you have thrown away in making those approximations.

We will use a simplified hybrid-π model (see for example [10]), in which the small-signal, i.e. AC component of the collector current, is given as $g_m v_{be}$ by a voltage-controlled current source, where v_{be} is the (small-signal) base-emitter voltage, developed across r_π, the resistance looking into the base. The transconductance g_m is that at the collector bias current of I_C, and is equal to I_C/V_T, which in our case comes down to $g_m = I_f/2V_T$.

Using a similar notational and numbering scheme as previously, though moving to lower case, the resultant models for the input, filter, and output stages are shown in Figure 6. As before we assume that the base currents are negligible, i.e. there are no currents through resistors r_π, so they could have been shown disconnected from the

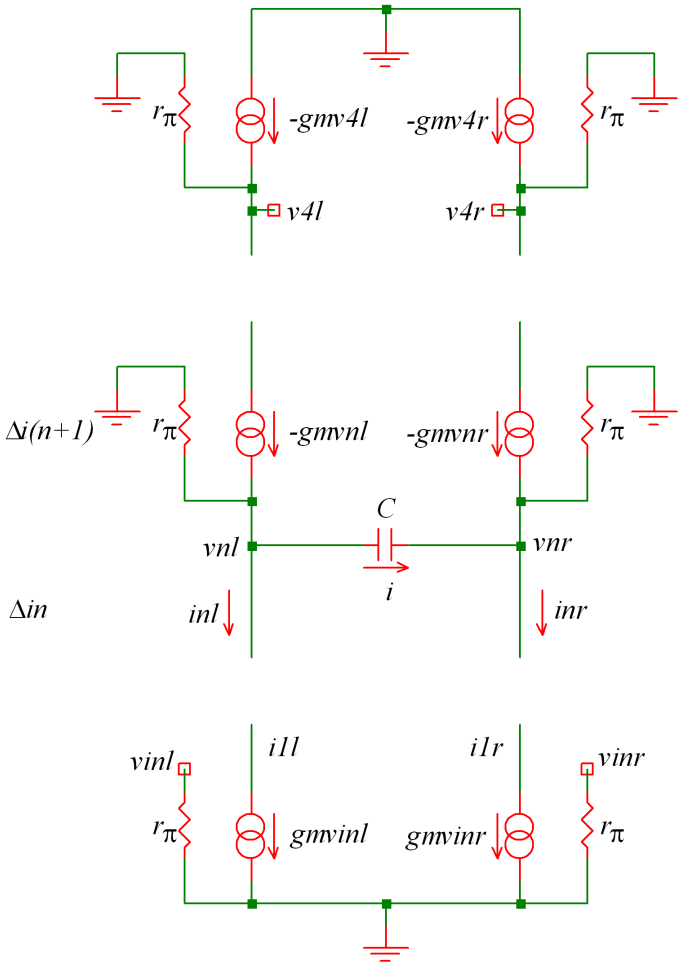


Figure 6: Small-signal models: input (bottom), filter stage (middle), and output (top)

emitter nodes: however I found it useful to leave them in, as it helps track the sign-sense of the voltages feeding the current sources.

The development closely mirrors that of the previous section, so I will not give all the details. Starting with the filter stage itself, looking at the currents we have

$$\begin{aligned} i_{(n+1)l} &= i_{nl} + i, \\ i + i_{(n+1)r} &= i_{nr} \end{aligned}$$

which subtract to give

$$\begin{aligned} i_{(n+1)l} - i_{(n+1)r} = \Delta i_{n+1} &= i_{nl} + i - (i_{nr} - i) = i_{nl} - i_{nr} + 2i \\ &= \Delta i_n + 2i. \end{aligned}$$

At the capacitor we have

$$v_{nl} - v_{nr} = \frac{i}{sC},$$

and by definition

$$i_{(n+1)l} = -g_m v_{nl} \quad \text{and} \quad i_{(n+1)r} = -g_m v_{nr},$$

so

$$\begin{aligned} i = (v_{nl} - v_{nr})sC &= -\frac{(i_{(n+1)l} - i_{(n+1)r})sC}{g_m} \\ &= -\frac{sC}{g_m}\Delta i_{n+1}, \end{aligned}$$

and on substituting for i in the expression above

$$\begin{aligned} \Delta i_{n+1} &= \Delta i_n + 2i \\ &= \Delta i_n - \frac{2sC}{g_m}\Delta i_{n+1}, \end{aligned}$$

gives the transfer function for stage n as

$$\frac{\Delta i_{n+1}}{\Delta i_n} = \frac{1}{\frac{2sC}{g_m} + 1}.$$

Substituting for g_m as

$$g_m = \frac{I_C}{V_T} = \frac{I_f}{2V_T}$$

then gives

$$\frac{\Delta i_{n+1}}{\Delta i_n} = \frac{1}{s\frac{4CV_T}{I_f} + 1}, \tag{18}$$

which is identical to equation (11) in the large-signal development above.

For the input pair we have by definition of the model:

$$i_{1l} = g_m v_{inl} \quad \text{and} \quad i_{1r} = g_m v_{inr},$$

so

$$\Delta i_1 = i_{1l} - i_{1r} = g_m(v_{inl} - v_{inr}) = g_m v_{in}$$

that is

$$\Delta i_1 = g_m v_{in}. \tag{19}$$

Similarly for the output pair:

$$i_{5l} = -g_m v_{4l} \quad \text{and} \quad i_{5r} = -g_m v_{4r},$$

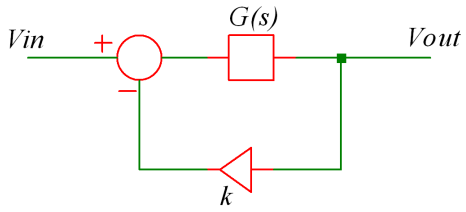


Figure 7: Block diagram of complete filter with feedback

and so

$$\Delta i_5 = i_{5l} - i_{5r} = -g_m(v_{4l} - v_{4r}) = -g_m v_{out}$$

that is

$$\Delta i_5 = -g_m v_{out}. \tag{20}$$

With equations (18), (19) and (20) the equivalent of (11), (14) and (15), it is easy to see that if we follow the same process starting at (16), we will indeed end up with the same transfer function for the filter core, equation (17).

2.3 Adding a feedback loop

The previous sections has been primarily concerned with the transfer function of the core of the filter: however, to make a versatile and interesting-sounding filter, any practical implementation will almost certainly have a variable-gain feedback path added in. How this is achieved at the level of the transfer function, and how the feedback affects the function is shown in this section.

Figure 7 is a simple block diagram of a complete filter, showing the transfer function of the filter core as $G(s)$, and with gain k feeding a proportion of the output V_{out} back around to the input. The output of the filter is then simply

$$V_{out} = G(s)(V_{in} - kV_{out}),$$

which is easily re-arranged to give

$$H(s) = \frac{V_{out}}{V_{in}} = \frac{G(s)}{1 + kG(s)}. \tag{21}$$

If we substitute the transfer function of the core as

$$G(s) = \frac{1}{(s+1)^4},$$

noting that this is the negation of equation (17) (a notational convenience, but easily effected by swapping the order of the output take-off at the top of the ladder), we get

$$H_{std}(s) = \frac{\dfrac{1}{(s+1)^4}}{1 + k\dfrac{1}{(s+1)^4}},$$

which on clearing the fractions gives

$$H_{std}(s) = \frac{1}{(s+1)^4 + k}, \tag{22}$$

or occasionally it is seen as

$$H_{std}(s) = \frac{1}{s^4 + 4s^3 + 6s^2 + 4s + 1 + k}$$

in expanded form (and where 'std' is an abbreviation of 'standard', for distinction with those coming later). This is the transfer function of a filter consisting of four identical, cascaded, buffered first-order sections, with an overall feedback gain amount k. Thus not only does it apply to the Moog ladder filter being considered here, it turns out that many other filters used in synthesizers share this transfer function, for example, many built using the CEM3320 and SSM2040 filter chips (and probably the SSM2044 too), and other circuits using four VCAs or OTAs, be they discrete or integrated—within the synthesizer domain, it is actually pretty hard to get away from this function if the filter is a 4-pole one (one notable exception being the *diode* ladders considered later).

2.4 Accuracy of the transistor ladder model

Before going on to examine the properties of the transfer function $H_{std}(s)$ (equation (22)) in more detail in the next section, it would seem prudent to check that it does model the filter reasonably well, at some sort of level. To this end some simple SPICE simulations were run in SIMetrix, and compared against data calculated from the transfer function using Mathematica—whilst the SPICE AC analyses used will make some of the same sorts of assumptions, for example linearizing the transistors about their bias points, it will act as a fairly 'gross' check that the model is able to predict the behaviour of the real circuit.

The simulation circuit used is shown in Figure 8. Much of the circuit has been 'idealized' so that we can concentrate on the effects of the filter core itself: the biasing resistors have been replaced with simple voltage sources, to ensure that transistors in the ladder are well separated, and with $V2$ providing an AC ground, so lifting the input pair off ground; feedback is provided by the voltage-controlled voltage source $E1$, whose output feeds through a (large) capacitor $C4$, which allows for the differing DC levels

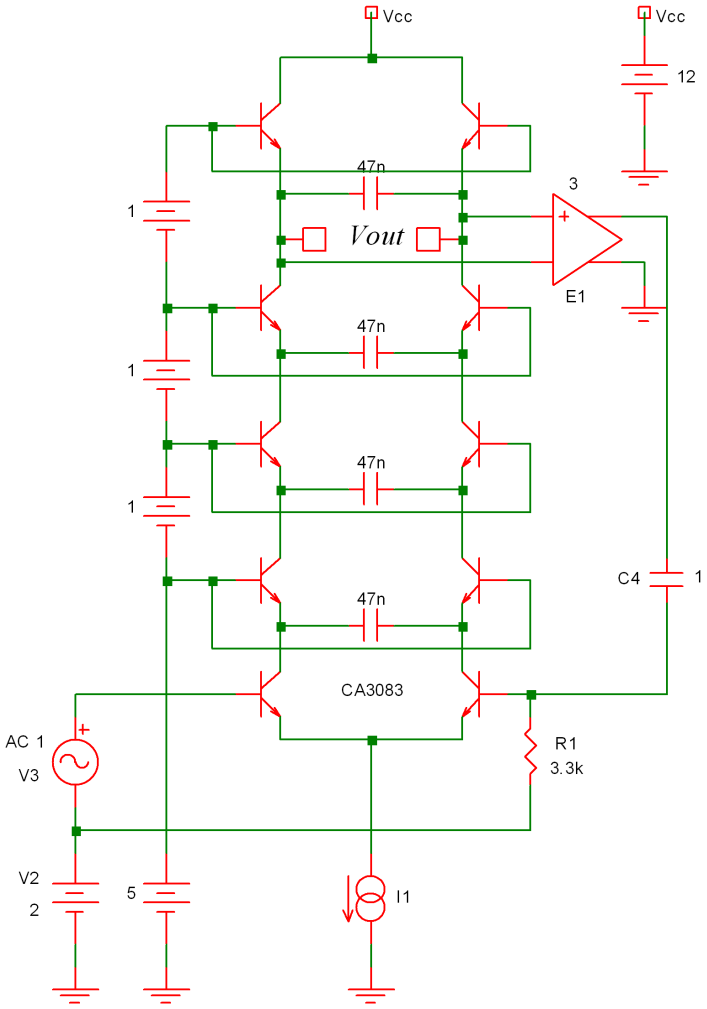


Figure 8: Simulation circuit for accuracy check

there, then onto $R1$, rather arbitrarily chosen to prevent the AC signal being 'grounded' at $V2$; DC current source $I1$ supplies the frequency-controlling current I_f; all transistors use the SPICE model for the CA3083, this being a practical contender for use in a real circuit; all capacitors are 47nF, calculated as giving practical cut-off frequencies for I_f ranging from around 10μA to 500μA. (Also note that no separate output transistor pair is used, the output being taken directly from top filter section, as was assumed in the analysis.)

Data was calculated in Mathematica by denormalizing $H_{std}(s)$, and switching from angular frequency ω to simply f in Hertz. Thus the following was calculated:

$$20 \log \left| H_{std} \left(\frac{4CV_T}{I_f} 2\pi f j, k \right) \right|,$$

for $k = 3$ and f ranging from 100Hz to 100kHz (and $V_T = 0.026$V, $C = 47$nF). The data was then exported from Mathematica and imported into SIMetrix, where it could

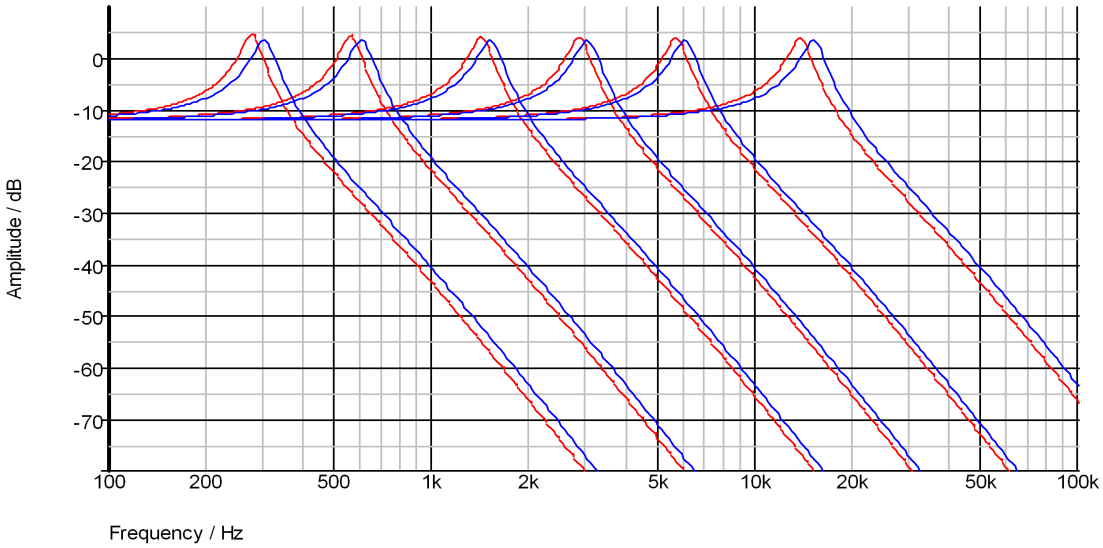


Figure 9: Plots from simulation of circuit, red, vs. calculation of H_{std}, blue

be plotted alongside the simulation results.

The results are shown in Figure 9, for six values of $I_f = 10, 20, 50, 100, 200, 500\mu$A, and with $k = 3$ we get quite a large amount of 'corner peaking'. These are generally in good agreement, but it is clear that the frequency of the calculations is slightly higher than the simulation—this seems to visually decrease a little with increasing I_f, to a minimum perhaps around $I_f = 100\mu$A, and then start increasing again. In Section 3.4 later we will see that the equivalent for the diode ladders shows a much closer agreement, leading me to think that this discrepancy is caused by the fact that the calculations assume all the base currents are zero, whereas in the simulation (and indeed a real circuit), some current will enter the ladder, and so have some impact on the DC biasing. Either way, at the very least equation (13) provides a simple calculation for determining the range cut-off frequencies in a design from the chosen C and range of I_f.

2.5 Poles and frequency responses

The poles of $H_{std}(s)$, i.e. the roots of the denominator, so solutions of $(1 + s)^4 + k = 0$, determine the shape of the frequency response, which is of some interest in the musical/synthesizer setting. One wouldn't normally expect to have the luxury of being able to determine the poles of a fourth-order filter analytically, and would probable opt for some numerical computation to determine them: however in this case the denominator is relatively simple, and so it is possible to solve for the poles. Taking

our cue from [4], proceed thus:

$$(s+1)^4 + k = 0$$

so

$$(s+1)^4 = -k$$

and then

$$s + 1 = (-k)^{1/4} = k^{1/4} e^{\frac{j\pi(2m+1)}{4}},$$

where we have utilized the odd 8^{th}-roots of unity, for $m = 0, 1, 2, 3$. Hence

$$s = -1 + k^{1/4} e^{\frac{j\pi(2m+1)}{4}},$$

which may be alternatively written as

$$s = -1 \pm k^{1/4} e^{\pm j\pi/4}.$$

From this it can be seen that the complex exponential term means the poles are evenly spaced in an 'X' pattern, at forty five degrees from the real axis, and the '-1' term shifts them all left, centering them about the point $(-1, 0)$. When $k = 0$, all four poles are *at* $(-1, 0)$; as k increases they move out along the arms of the 'X', equidistant from the point $(-1, 0)$; at $k = 4$, the real component of the rightmost pair is zero, i.e. they reach the imaginary axis, and the filter will be oscillating. Figure 10 shows the location of poles on the complex s-plane ($s = \sigma + \omega j$), and how they migrate for about 20 values of k as it is varied from 0 to 4[†] (the offset axis scaling is for uniformity of comparison with the plots of the diode filters yet to come).

The frequency responses of the filter (both magnitude and phase) are of course essentially defined by the pole locations. The magnitude response is defined as $|H_{std}(\omega j)|$ or more usually as $20 \log |H_{std}(\omega j)|$, which is itself the intersection of the 2-dimensional surface $20 \log |H_{std}(s)|$ with the plane which goes through the imaginary axis, and which is perpendicular to the s-plane itself (set the real component, σ, in $s = \sigma + \omega j$ to zero). It is a relatively simple matter to draw such surfaces using a package such as 'Mathematica': Figure 11 shows the surface for $H_{std}(s) = 1/((s+1)^4 + k)$ at $k = 0.5$. The peaks in the surface are caused by the poles, which even though k is relatively small still at 0.5, are already quite far apart from each other; the 'slice' through the surface

[†]To get the even spacing, k was varied as $4(n/20)^4$ for $n = 0, 1, \ldots 20$, so the actual values of k are: 0.0, 0.000025, 0.0004, 0.002025, 0.0064, 0.015625, 0.0324, 0.060025, 0.1024, 0.164025, 0.25, 0.366025, 0.5184, 0.714025, 0.9604, 1.26563, 1.6384, 2.08803, 2.6244, 3.25803, 4.0.

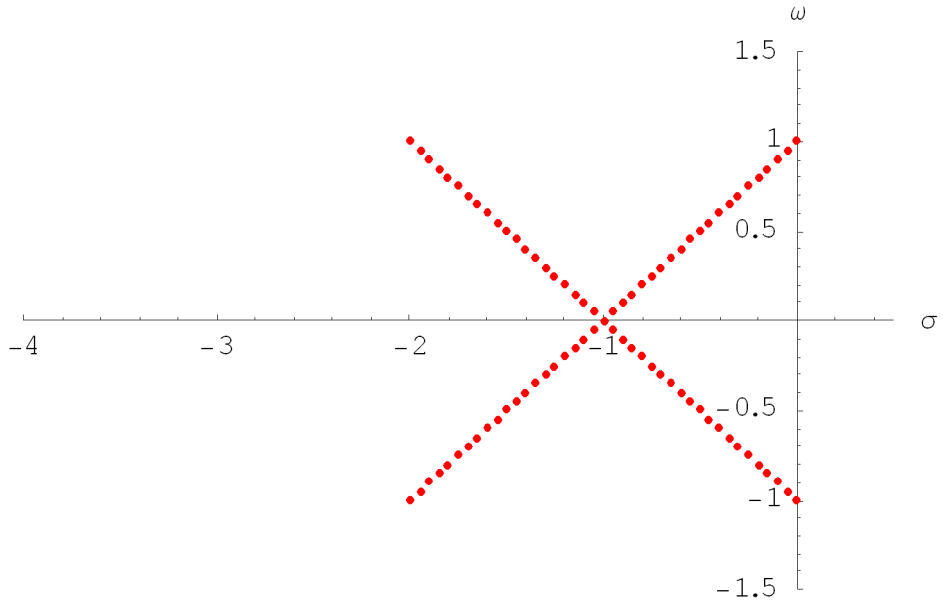


Figure 10: Poles of $1/((s+1)^4+k)$, $k=0$ to 4

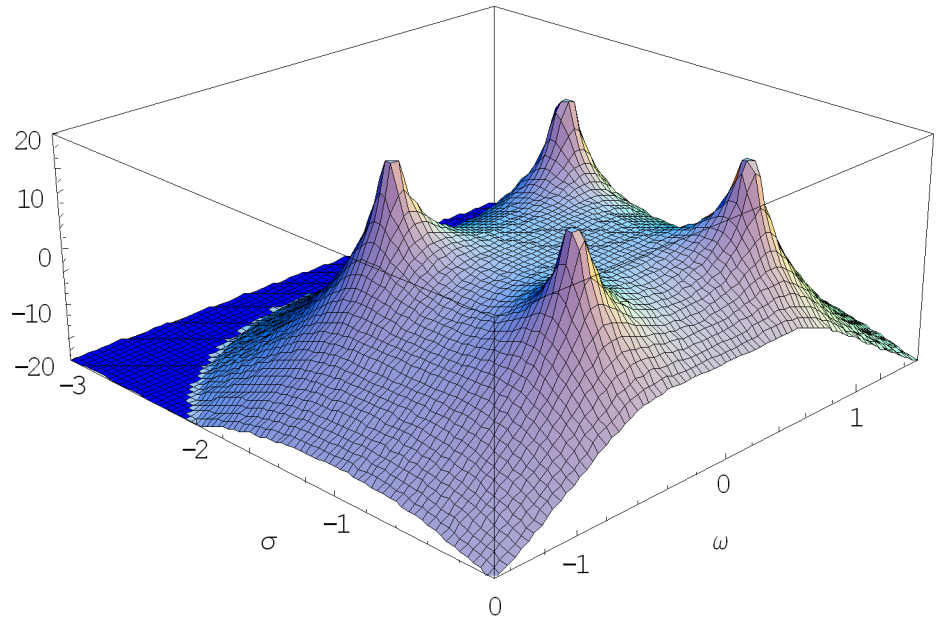


Figure 11: $20\log|H_{std}(s)|$ for $k=0.5$

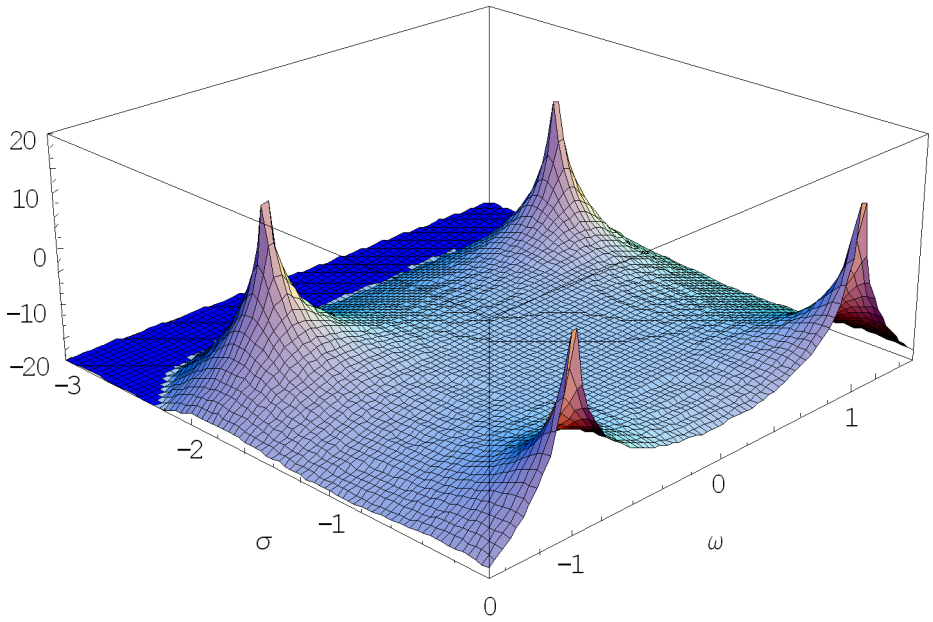


Figure 12: $20\log|H_{std}(s)|$ for $k = 4$

at the right-hand edge is the normal frequency response curve, $20\log|H_{std}(\omega j)|$, but note that the scaling along the imaginary axis, ω, is linear, and not the more normal logarithmic one used for simple 2-D plots (which follow below). As k is still small, the rightmost pair of poles are still some way from the imaginary axis, and it can be seen that this results in a passband which is still quite flat, unlike in the following figure. In the self-resonating condition, with $k = 4$, Figure 12 shows that the rightmost pair of poles have now reached the imaginary axis, giving the large resonant peak in the response, and because the poles are now so far apart, the surface is able to 'sag' between them, producing a considerable 'droop' in the passband.[†]

More traditional 'Bode plot' frequency responses are shown in Figure 13, for values of k of $0, 0.5, 1, 1.5, 2, 2.5, 3, 4$: $k = 0$ is the least 'peaky', at 0dB where it meets the y-axis; $k = 4$ is the biggest peak, at approximately -14dB at the y-axis. These curves immediately highlight several issues worthy of note. First, the 'droop' in the passband as the resonance is increased is seen by many as big a failing of this type of filter, and this feature provides an easy test to see if a filter is of this type: open the filter right up, by setting a high cut-off frequency; input a signal at a frequency well inside the passband; turn the resonance up, and if the level of the signal drops, chances are the filter is of this type (and note diode ladders suffer similarly).

Secondly I think it highlights something of a 'terminological issue', which whilst it

[†]Animated gifs showing the smooth transition of the surface as the poles move due to varying k may be viewed at my website: `http://www.timstinchcombe.co.uk/synth/poles/poles.html`

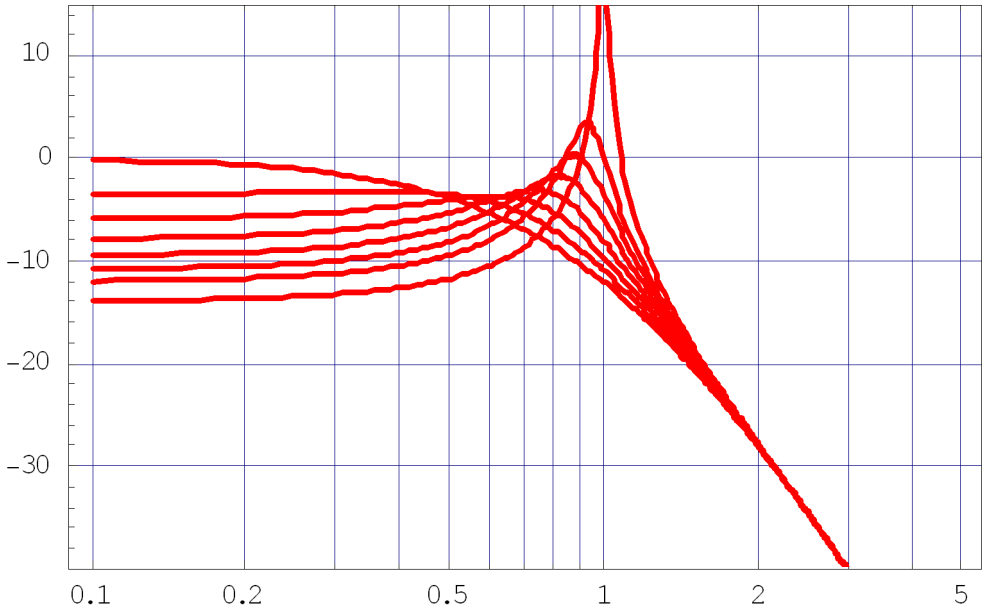


Figure 13: Magnitude responses of $1/((s+1)^4+k)$, various k

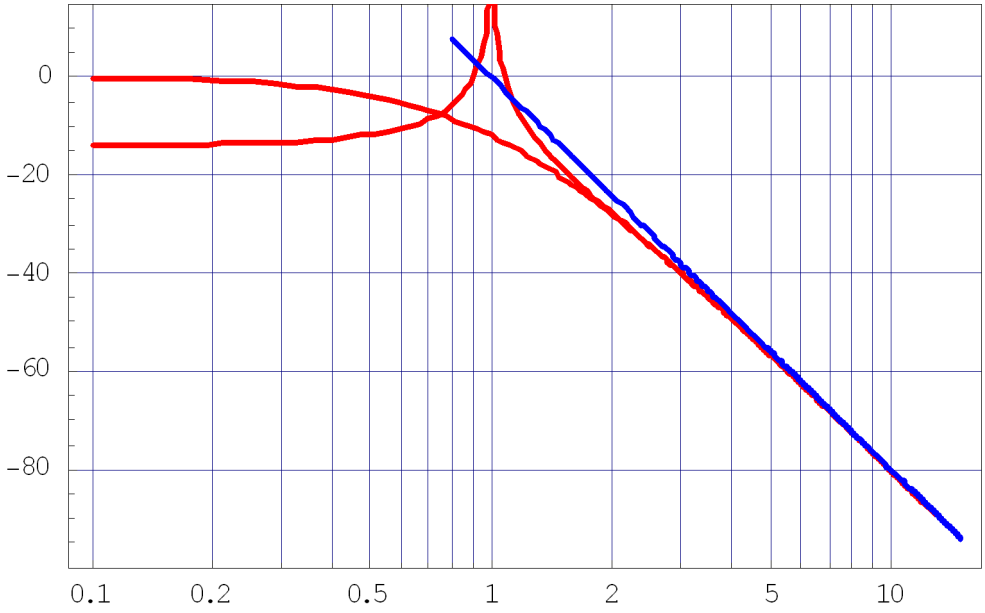


Figure 14: Magnitude responses of $H_{std}(s)$ at $k=0$ and 4, and the asymptote '$1/\omega^4$'

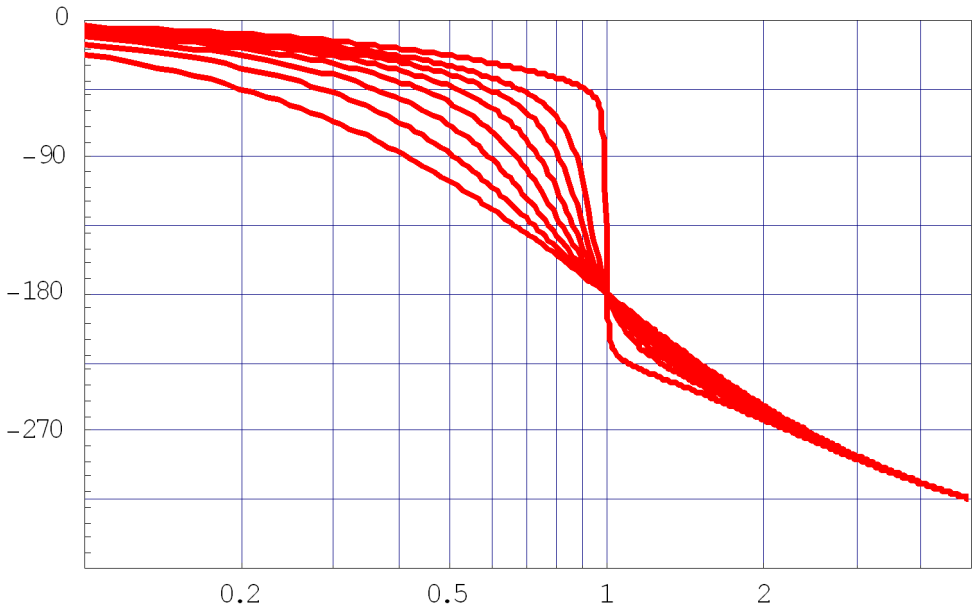


Figure 15: Phase responses of $1/((s+1)^4+k)$, various k

my not have the direst of consequences, certainly adds a lot of 'fog' when discussing these filters: what point in all these curves do we actually mean when we talk of the *cut-off frequency* of the filter? The first-order filter concept of the '3dB' or 'half-power' point doesn't relate very well to higher-order cases once we introduce resonance into the proceedings: certainly the $k = 0$ (flattest) curve in Figure 13 cuts $\omega = 1$ at -12dB (being 4×-3dB), but for $k > 0$, choosing ω at the point where the curve crosses -12dB would result in all manner of differing ω values. Alternatively, choosing ω at the point where the peak occurs will also clearly result in ω values which wander all over the place. As unsatisfactory as it may seem, it seems we have to accept that all the curves 'belong' to $\omega = 1$ (reminding ourselves we are dealing with a normalized function here), on the grounds that the asymptote to the curves in the stopband passes through 0dB *at* $\omega = 1$ (which is equivalent to the '3dB' rule multiplied up). (Visually trying to align the asymptote to the curves in Figure 13 suggests this might not be the case, but if we look further down into the stopband where the curves have stopped 'bending round', the true picture *is* seen, as shown in Figure 14 where the asmyptote is drawn along with two of the curves.) Of course to get the peaks to 'line up' better etc. requires a different transfer function, i.e. we need a different filter!

Some phase responses of the function $H_{std}(s)$ are shown in Figure 15, for the same values of k as used previously for the magnitude responses, except for the largest value which is 3.9 and not 4 (the 180° 'flip' at $\omega = 1$ for $k = 4$ is so severe it looks as though the plot has been 'doctored' in some way!). The $k = 0$ curve is the flattest, each curve becoming more contorted as k is increased. I currently have little idea as to how

unremarkable or otherwise these plots might be.

3 Diode Ladder Filters

3.1 Introduction

Replacing the transistors of the Moog ladder filter with diodes results in a four-pole filter which shares some similarities with the original filter, but which also has a lot of differences. The main difference is probably the fact that each section is no longer buffered from the next—this means that all of the relative simplicity of the original is lost, resulting in a much less elegant looking transfer function (and which is considerably harder to derive!).

In the early days, using a diode ladder filter was probably done to avoid infringing Moog's patent on the original filter ([2]): using diodes actually adds a little flexibility, and so quite a lot of different configurations appear to have sprung up. The basic core designs of some of these are shown in Figures 16 and 17, and they show a number of features worthy of comment.

Firstly we note that there are a varying number of diodes at the top of the ladder: with one diode, the Roland TB-303 (OK so it is a transistor, but analytically it is the same as if it were a diode!); with two, the Practical Electronics Minisonic and the Roland 100; and with three, the EMS filters and the Doepfer A-102. Because of the way these diodes set the bias voltage down the ladder, using more of them increases the basic gain of the filter, as will be seen below, and it also has a small impact on the frequency response too.

Some of the designers have chosen to use a smaller capacitor at the bottom of the ladder, typically half the value of the others (TB-303, Doepfer A-102 and Roland 100). As of this writing I am not completely sure of the purpose of this—certainly we will see below that it has minimal impact on the pole locations. (I suspect it might be for some stability issue which I don't currently comprehend.)

Several of the designs have extraneous diodes at the bottom of the ladder, between the differential input pair and the first filter section (Doepfer A-102, early EMS, PE Minisonic): these do not contribute to the filtering action, and so appear to be quite redundant.

Two of the designs are actually *five*-pole filters—the later EMS and the Roland 100. I didn't become aware of them until some time after I had done the bulk of the analysis in this paper, so they are not covered in detail. (The methods used below though are fairly general, and so it should be a relatively easy matter to extend them to cover these filters: thus their analysis is 'left as an exercise for the reader' as they say—perhaps

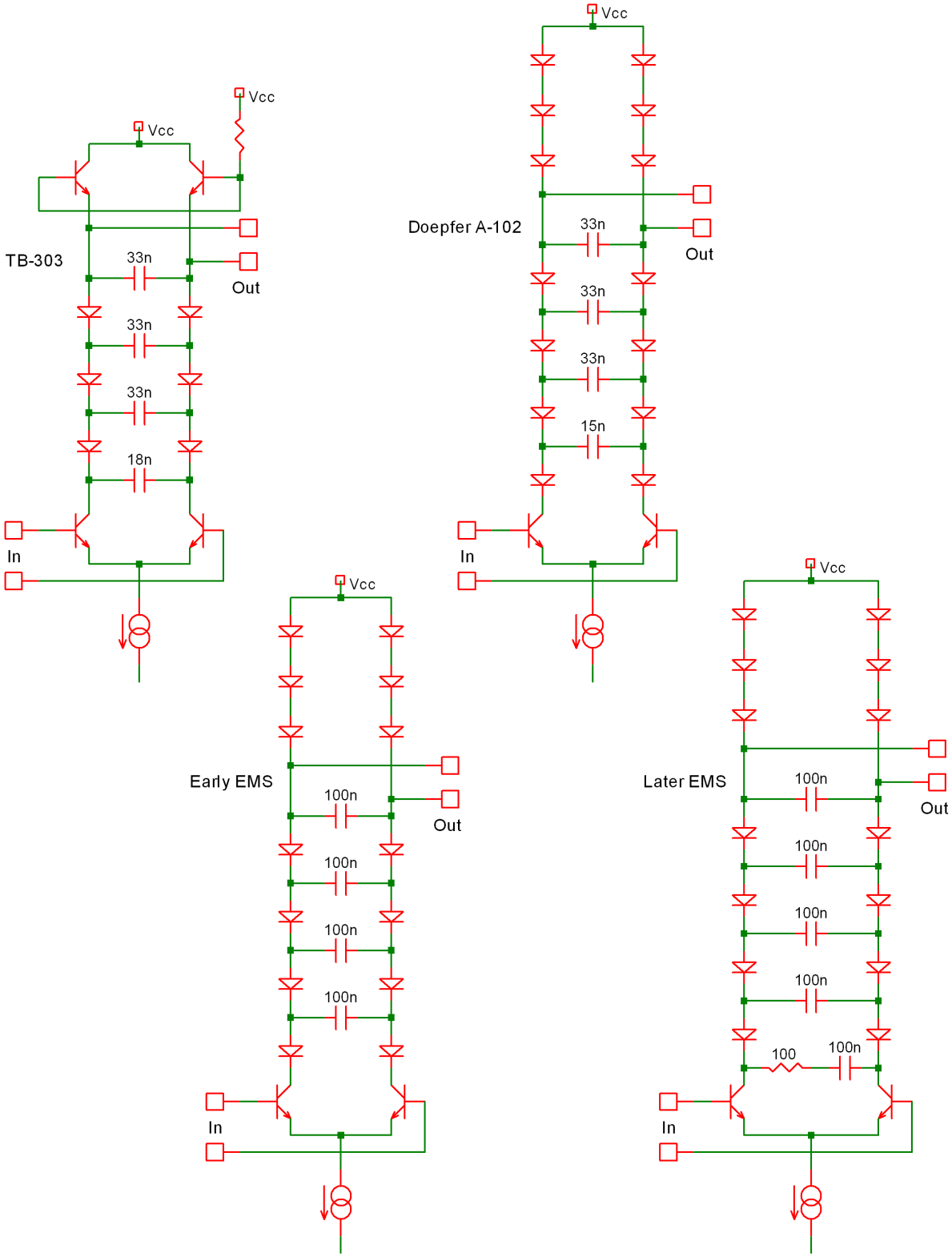


Figure 16: Various diode ladder filter cores

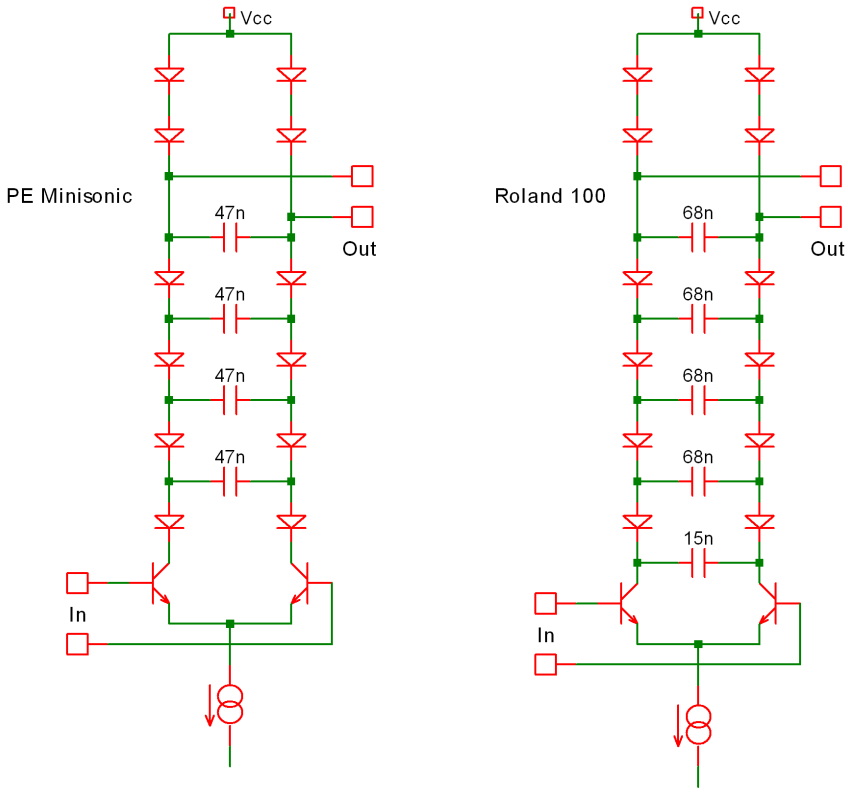


Figure 17: More diode ladder filter cores

some aspiring student might be willing to 'fill in the blanks'?!) The 100Ω resistor in the later EMS design actually introduces a zero into the transfer function, and we return briefly to look at this in Section 3.6 later.

On seeing a degree of this variation in these circuits when I originally started out on this analysis several years ago, I decided to make it reasonably general, so that it would cover many of the different cases. To that end I allowed for a differing number of diodes at the top of the ladder, and for the lower capacitor to be different from the rest. However I was at that time working with a Doepfer A-102 as a then-believed copy of the EMS filter, and was unaware of the five-pole types. I have now seen some EMS VCS3 schematics, and also have the Analogue Systems RS500e 'authorized copy' of the EMS filter, and have discovered that the A-102 doesn't follow the EMS filter exactly, particularly the capacitor ratios (but then in fairness to Doepfer I'm not sure that in the end they intended it to be an exact copy anyway). Thus what I originally thought to be the 'EMS configuration' of 3 diodes and the bottom capacitor halved turned out not to be the case, and the A-102 is thus the only example I have currently seen which is like this. However I decide to leave it in anyway, as it illustrates a point. There is also a copy of an anonymous schematic (which is why I feel it worthy of mention) of another EMS-like filter (3 diodes; equal caps) floating around on the Internet: this is

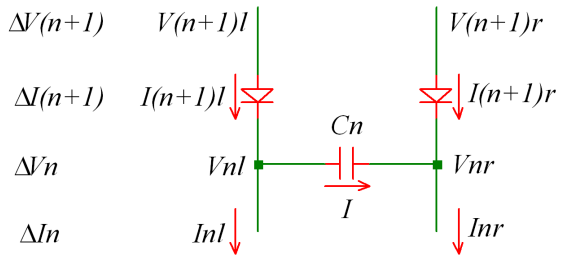


Figure 18: Notation for diode ladder stage n

in fact from [11], published by Babani, whose series of books has a very recognisable print-style once you know it!

The following section therefore works the transfer function in a semi-general sense, deriving four cases in particular:

- one diode and equal capacitors: the simplest case. Done mainly as it allows a cross-check against the results of others who have derived this transfer function (though I don't currently know of an actual example of this configuration, but someone somewhere has probably built one!)

- one diode with the bottom capacitor halved: as the TB-303

- 3 diodes and equal capacitors: as the EMS filters

- 3 diodes and the bottom capacitor halved: as the A-102

The PE Minisonic configuration of 2 diodes and equal capacitors should be easily deducible from these results, and as stated earlier, it shouldn't be too hard to extend them to cover the 5-pole filters too!

3.2 Transfer function derivations

In the transistor ladder, apart from the differential voltage across the capacitor aiding the filtering action, the voltages *down* the ladder don't really have a part to play, as they are kept separated from each other by the chain of biasing resistors. In the diode ladder this is a different matter though, as the top of the ladder sets the reference point, and *all* voltages down the ladder derive from it. Thus in order to work the transfer function out, all intermediate voltages need to be eliminated, and this appears to be quite a hurdle to overcome. My solution to this problem is to define the differences in voltage and current at each stage *in terms of the stage above*: one can then start at the bottom and work up, eliminating all the intermediate values as you go. Figure 18 shows the notation used about stage n, and which is similar to that used previously: currents

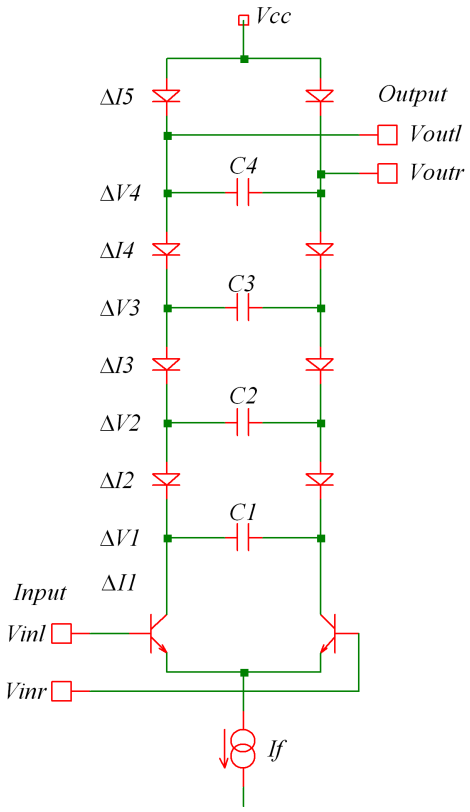


Figure 19: Basic diode ladder core and stage numbering

in the ladder arms *below* the stage capacitor C_n are I_{nl} and I_{nr}, and so their difference is

$$\Delta I_n = I_{nl} - I_{nr}.$$

The voltages either side of the capacitor are V_{nl} and V_{nr}, with difference

$$\Delta V_n = V_{nl} - V_{nr},$$

and the currents *above* the capacitor are $I_{(n+1)l}$ and $I_{(n+1)r}$, so

$$\Delta I_{n+1} = I_{(n+1)l} - I_{(n+1)r}.$$

The full numbering of the stages against the basic core of the filter we are going to use is shown in Figure 19: also note that, as the transistor case, we are drawing current I_f, which sets the cut-off frequency, out of the bottom of the ladder—and since no current can enter or leave the ladder except at the bottom or top, we must have

$$I_{nl} + I_{nr} = I_f$$

for all n.

Now, from Figure 18 we have

$$\begin{aligned} I_{(n+1)l} &= I + I_{nl}, \\ I_{(n+1)r} + I &= I_{nr}, \end{aligned}$$

and subtracting gives

$$I_{(n+1)l} - I_{(n+1)r} = \Delta I_{n+1} = I + I_{nl} + I - I_{nr} = \Delta I_n + 2I$$

so

$$\Delta I_{n+1} = \Delta I_n + 2I.$$

Also at the capacitor we get

$$V_{nl} - V_{nr} = \Delta V_n = \frac{I}{sC_n}$$

so ridding I above gives

$$\Delta I_{n+1} = \Delta I_n + 2sC_n\Delta V_n,$$

and re-arranging this gives

$$\Delta I_n = \Delta I_{n+1} - 2sC_n\Delta V_n. \tag{23}$$

We now seek to rid the stage n term, ΔV_n, from the right-hand side of equation (23). If we use the standard diode equation, and write down the ratio of the $(n+1)$-stage currents, we get

$$\begin{aligned} \frac{I_{(n+1)r}}{I_{(n+1)l}} &= \frac{e^{\frac{V_{(n+1)r} - V_{nr}}{V_T}}}{e^{\frac{V_{(n+1)l} - V_{nl}}{V_T}}} = e^{\frac{V_{(n+1)r} - V_{(n+1)l} - (V_{nr} - V_{nl})}{V_T}} \\ &= e^{\frac{-\Delta V_{n+1} + \Delta V_n}{V_T}}, \end{aligned}$$

and as already noted, the currents sum to I_f, so we can use equation (9) to immediately get

$$\begin{aligned} \Delta I_{n+1} &= \frac{I_f}{2V_T}(\Delta V_{n+1} - \Delta V_n) \\ &= \frac{1}{a}(\Delta V_{n+1} - \Delta V_n), \quad \text{say, where} \quad a = \frac{2V_T}{I_f}, \end{aligned}$$

or

$$\Delta V_n = \Delta V_{n+1} - a\Delta I_{n+1}. \tag{24}$$

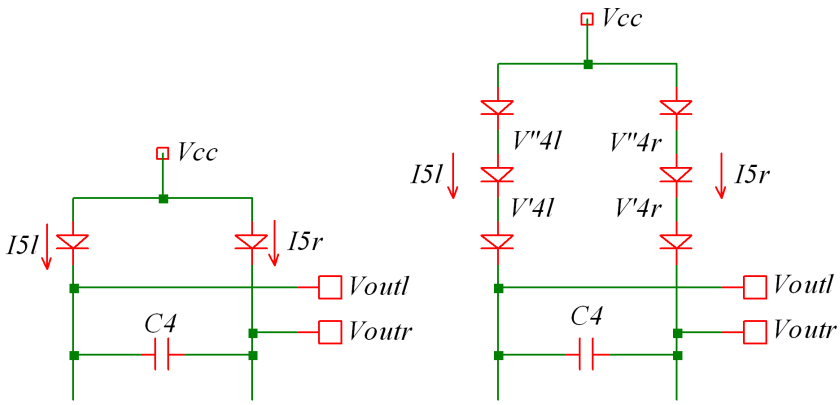


Figure 20: Output stage notation for one and three diodes

(The symbol a is just introduced for notational convenience, in order to help reduce the number of symbols that require tracking in the following working.) Now substitute for ΔV_n here into equation (23):

$$\Delta I_n = \Delta I_{n+1} - 2sC_n(\Delta V_{n+1} - a\Delta I_{n+1})$$

which re-orders as

$$\Delta I_n = (1 + 2asC_n)\Delta I_{n+1} - 2sC_n\Delta V_{n+1}. \tag{25}$$

Equations (24) and (25) only contain voltage and current terms for stage '$n+1$' in their right-hand sides, and so we can now use them to work up the ladder. However before we do so, it will help to sort out the input and output stages.

The input stage is easy—it is the same as the transistor ladder, so it is equation (14):

$$\Delta I_1 = \frac{I_f V_{in}}{2V_T} = \frac{V_{in}}{a}.$$

The output pair is only a little more involved: work directly from the diode equations for a single diode at the top, and then do three, from which the general pattern for any number of diodes is easily seen. Referring to Figure 20, from the standard diode equation we have

$$I_{5l} = I_s e^{\frac{V_{cc}-V_{outl}}{V_T}} \qquad \text{and} \qquad I_{5r} = I_s e^{\frac{V_{cc}-V_{outr}}{V_T}}. \tag{26}$$

(For the two transistors in the TB-303 case, you get almost identical expressions, the only difference being that the reference voltage is that on the base, rather than the bias voltage that I've called V_{cc} here, and in any case it disappears in the next step!) Dividing these two gives

$$\frac{I_{5r}}{I_{5l}} = e^{\frac{V_{outl}-V_{outr}}{V_T}}.$$

Again as $I_f = I_{5l} + I_{5r}$ (recall no current into or out of the ladder), this time we call on (8) to get:

$$\Delta I_5 = I_{5l} - I_{5r} = -\frac{I_f(V_{outl} - V_{outr})}{2V_T} = -\frac{V_{out}}{a}, \tag{27}$$

noting the subtle sign change (the common voltage is at the top and not the bottom!). Next, for three diodes, write down the expressions for each diode on each side:

$$\begin{aligned}
I_{5l} &= I_s e^{\frac{V_{cc} - V''_{4l}}{V_T}} \\
I_{5l} &= I_s e^{\frac{V''_{4l} - V'_{4l}}{V_T}} \\
I_{5l} &= I_s e^{\frac{V'_{4l} - V_{outl}}{V_T}} \\
I_{5r} &= I_s e^{\frac{V_{cc} - V''_{4r}}{V_T}} \\
I_{5r} &= I_s e^{\frac{V''_{4r} - V'_{4r}}{V_T}} \\
I_{5r} &= I_s e^{\frac{V'_{4r} - V_{outr}}{V_T}},
\end{aligned}$$

and then multiply down each side for the left and right expressions, and divide the right by the left:

$$\frac{I_{5r}^3}{I_{5l}^3} = \frac{e^{\frac{V_{cc} - V''_{4r}}{V_T}} e^{\frac{V''_{4r} - V'_{4r}}{V_T}} e^{\frac{V'_{4r} - V_{outr}}{V_T}}}{e^{\frac{V_{cc} - V''_{4l}}{V_T}} e^{\frac{V''_{4l} - V'_{4l}}{V_T}} e^{\frac{V'_{4l} - V_{outl}}{V_T}}} = e^{\frac{V_{outl} - V_{outr}}{V_T}},$$

so

$$\frac{I_{5r}}{I_{5l}} = e^{\frac{V_{outl} - V_{outr}}{3V_T}}$$

Again conditions of equation (8) are met, but this time we need to modify the expression using '$3V_T$', and not 'V_T', thus giving

$$\Delta I_5 = I_{5l} - I_{5r} = -\frac{I_f(V_{outl} - V_{outr})}{2 \cdot 3V_T} = -\frac{V_{out}}{3a}. \tag{28}$$

Looking back at equation (27) it is clear to generalize to d diodes we need only write

$$\Delta I_5 = -\frac{V_{out}}{da}. \tag{29}$$

(And we should also check that the conditions for the linear approximation used in the derivation of (7) are still met—they will be, as the division by $d > 1$ actually makes the tanh argument *smaller*, which helps.)

Now we can return to the business of recursively working up the ladder. Start with the expression for the input

$$\frac{V_{in}}{a} = \Delta I_1,$$

and use equation (25) with $n = 1$:

$$\frac{V_{in}}{a} = \Delta I_1 = (1 + 2asC_1)\Delta I_2 - 2sC_1\Delta V_2.$$

Some more notational pruning seems prudent: assume that only the first capacitor is different from the rest, so write $C_n = C$ for $n \geqslant 2$. Now rid ΔI_2 and ΔV_2 by using (25) and (24) with $n = 2$,

$$\begin{aligned}\frac{V_{in}}{a} &= (1 + 2asC_1)\big(\Delta I_3(1 + 2asC) - 2sC\Delta V_3\big) - 2sC_1(\Delta V_3 - a\Delta I_3)\\ &= \big(4a^2s^2C_1C + 2as(2C_1 + C) + 1\big)\Delta I_3 - \big(4as^2C_1C + 2s(C_1 + C)\big)\Delta V_3,\end{aligned}$$

and again for ΔI_3 and ΔV_3, $n = 3$:

$$\begin{aligned}&= \big(4a^2s^2C_1C + 2as(2C_1 + C) + 1\big)\big(\Delta I_4(1 + 2asC) - 2sC\Delta V_4\big)\\ &\quad - \big(4as^2C_1C + 2s(C_1 + C)\big)(\Delta V_4 - a\Delta I_4)\\ &= \big(8a^3s^3C_1C^2 + 4a^2s^2(4C_1C + C^2) + 6as(C_1 + C) + 1\big)\Delta I_4\\ &\quad - \big(8a^2s^3C_1C^2 + 4as^2(3C_1C + C^2) + 2s(C_1 + 2C)\big)\Delta V_4,\end{aligned}$$

and this time just ΔI_4, from (23) with $n = 4$:

$$\begin{aligned}&= \big(8a^3s^3C_1C^2 + 4a^2s^2(4C_1C + C^2) + 6as(C_1 + C) + 1\big)(\Delta I_5 - 2sC\Delta V_4)\\ &\quad - \big(8a^2s^3C_1C^2 + 4as^2(3C_1C + C^2) + 2s(C_1 + 2C)\big)\Delta V_4\\ &= \big(8a^3s^3C_1C^2 + 4a^2s^2(4C_1C + C^2) + 6as(C_1 + C) + 1\big)\Delta I_5\\ &\quad - \big(16a^3s^4C_1C^3 + 8a^2s^3(5C_1C^2 + C^3) + 4as^2(6C_1C + 4C^2) + 2s(C_1 + 3C)\big)\Delta V_4,\end{aligned}$$

and finally (29) for ΔI_5, and $\Delta V_4 = V_{4l} - V_{4r} \equiv V_{outl} - V_{outr} = V_{out}$:

$$\begin{aligned}&= -\big(8a^3s^3C_1C^2 + 4a^2s^2(4C_1C + C^2) + 6as(C_1 + C) + 1\big)\frac{V_{out}}{da}\\ &\quad - \big(16a^3s^4C_1C^3 + 8a^2s^3(5C_1C^2 + C^3) + 4as^2(6C_1C + 4C^2) + 2s(C_1 + 3C)\big)V_{out}.\end{aligned}$$

After a final gathering of terms, and leaving the expression up the wrong way, so as to avoid having to split the long fraction across the line break, we arrive at a (fairly) general (inverted) transfer function of:

$$\begin{aligned}-d\frac{V_{in}}{V_{out}} &= 8a^3s^3C_1C^2 + 4a^2s^2(4C_1C + C^2) + 6as(C_1 + C) + 1\\ &\quad + d\big(16a^4s^4C_1C^3 + 8a^3s^3(5C_1C^2 + C^3) + 4a^2s^2(6C_1C + 4C^2) + 2as(C_1 + 3C)\big),\end{aligned}$$

so

$$\begin{aligned}\frac{-d}{G(s)} &= 16a^4s^4dC_1C^3 + 8a^3s^3\big(C_1C^2(1 + 5d) + dC^3\big)\\ &\quad + 4a^2s^2\big(C_1C(4 + 6d) + C^2(1 + 4d)\big) + 2as\big(C_1(3 + d) + C(3 + 3d)\big) + 1.\end{aligned} \tag{30}$$

We can now substitute for C_1 and d into this and arrive at the transfer functions for the different configurations. First take the simplest case, one diode at the top, so $d = 1$, and all capacitors equal, so $C_1 = C$. Re-arrange to get a slightly more familiar form:

$$G_1(s) = \frac{-1}{16a^4s^4C^4 + 56a^3s^3C^3 + 60a^2s^2C^2 + 20asC + 1}.$$

Now put

$$\omega_c^4 = \frac{1}{16a^4C^4},$$

so that

$$\omega_c = \frac{1}{2aC} = \frac{I_f}{4CV_T},$$

from which we have

$$f_c = \frac{I_f}{8\pi CV_T},$$

to get

$$G_1(s) = \frac{-1}{\frac{s^4}{\omega_c^4} + 7\frac{s^3}{\omega_c^3} + 15\frac{s^2}{\omega_c^2} + 10\frac{s}{\omega_c} + 1},$$

and finally normalize the frequency to get

$$G_1(s) = \frac{-1}{s^4 + 7s^3 + 15s^2 + 10s + 1}. \tag{31}$$

This result has been reported by others, so the fact that we have ended up with it here is hopefully a partial validation that all the above algebraic wrangling is not entirely incorrect! We will examine it in detail later, after we have worked some of the other configurations.

For the TB-303 configuration of a single diode and the bottom capacitor half the others, put $d = 1$ and $C_1 = C/2$ into equation (30):

$$G_{tb}(s) = \frac{-1}{8a^4s^4C^4 + 32a^3s^3C^3 + 40a^2s^2C^2 + 16asC + 1}.$$

So this time

$$\omega_c^4 = \frac{1}{8a^4C^4}, \qquad \text{giving} \qquad \omega_c = \frac{1}{2^{\frac{3}{4}}aC} = \frac{I_f}{2^{\frac{7}{4}}CV_T},$$

so that

$$G_{tb}(s) = \frac{-1}{\frac{s^4}{\omega_c^4} + 2^{\frac{11}{4}}\frac{s^3}{\omega_c^3} + 10\sqrt{2}\frac{s^2}{\omega_c^2} + 2^{\frac{13}{4}}\frac{s}{\omega_c} + 1},$$

and finally normalize and evaluate constants to get

$$G_{tb}(s) = \frac{-1}{s^4 + 6.727s^3 + 14.142s^2 + 9.514s + 1}.$$

Next, for the EMS-type configuration with 3 diodes and equal capacitors, i.e. $d = 3$, $C_1 = C$:

$$G_{ems}(s) = \frac{-3}{48a^4s^4C^4 + 152a^3s^3C^3 + 140a^2s^2C^2 + 36asC + 1},$$

and this time

$$\omega_c^4 = \frac{1}{48a^4C^4}, \qquad \text{giving} \qquad \omega_c = \frac{1}{2 \cdot 3^{\frac{1}{4}}aC} = \frac{I_f}{3^{\frac{1}{4}}4CV_T}, \tag{32}$$

so that

$$G_{ems}(s) = \frac{-3}{\frac{s^4}{\omega_c^4} + \frac{19}{3^{\frac{3}{4}}}\frac{s^3}{\omega_c^3} + \frac{35}{\sqrt{3}}\frac{s^2}{\omega_c^2} + 2 \cdot 3^{\frac{7}{4}}\frac{s}{\omega_c} + 1},$$

finally giving

$$G_{ems}(s) = \frac{-3}{s^4 + 8.335s^3 + 20.207s^2 + 13.677s + 1}$$

on evaluating the constants.

Finally for the Doepfer A-102 configuration of 3 diodes and lower capacitor half the others, so $d = 3$ and $C_1 = C/2$:

$$G_d(s) = \frac{-3}{24a^4s^4C^4 + 88a^3s^3C^3 + 96a^2s^2C^2 + 30asC + 1},$$

so now

$$\omega_c^4 = \frac{1}{24a^4C^4}, \qquad \text{giving} \qquad \omega_c = \frac{1}{2^{\frac{3}{4}}3^{\frac{1}{4}}aC} = \frac{I_f}{2^{\frac{7}{4}}3^{\frac{1}{4}}CV_T}, \tag{33}$$

which in turn gives

$$G_d(s) = \frac{-3}{\frac{s^4}{\omega_c^4} + \frac{2^{\frac{3}{4}}11}{3^{\frac{3}{4}}}\frac{s^3}{\omega_c^3} + 2^{\frac{7}{2}}\sqrt{3}\frac{s^2}{\omega_c^2} + 2^{\frac{1}{4}}3^{\frac{3}{4}}5\frac{s}{\omega_c} + 1},$$

and finally

$$G_d(s) = \frac{-3}{s^4 + 8.116s^3 + 19.596s^2 + 13.554s + 1}.$$

3.3 Adding the feedback loops

In order to examine the effects that feedback has on each of the functions in the previous section, we need to put each $G_{(\cdot)}(s)$ into equation (21) in turn. In doing so I have again

switched the sign in the numerators, in order to avoid the negative sign on the constant in the denominator, to get the following:

$$\begin{aligned} H_1(s) &= \frac{1}{s^4 + 7s^3 + 15s^2 + 10s + 1 + k}, \\ H_{tb}(s) &= \frac{1}{s^4 + 6.727s^3 + 14.142s^2 + 9.514s + 1 + k}, \\ H_{ems}(s) &= \frac{3}{s^4 + 8.335s^3 + 20.207s^2 + 13.677s + 1 + 3k}, \\ H_d(s) &= \frac{3}{s^4 + 8.116s^3 + 19.596s^2 + 13.554s + 1 + 3k}. \end{aligned} \tag{34}$$

Various features and facets of these functions will now be examined in the following few sections.

3.4 Accuracy of the diode ladder models

As we did for the Moog transistor ladder, we now check some of these transfer functions against a simulation, to ensure that they do model the filters they represent reasonably well. First we do a TB-303-type filter, with the lower capacitor (literally) half the value of the others, remaining at 47nF as before. The circuit used is shown in Figure 21, and similar comments apply to the idealized elements as before. Of particular note is that all the diodes are diode-connected CA3083 transistors—the significance of this will be seen in a minute. De-normalizing expression (34) for the calculations gives:

$$20 \log \left| H_{tb}\left(\frac{2^{\frac{7}{4}} C V_T}{I_f} 2\pi f j, k \right) \right|,$$

and this time k needs to be slightly more for a similar amount of resonance, and was taken as $k = 10$ for the plots. As before f ranges from 100Hz to 100kHz, and I_f was stepped through $10, 20, 50, 100, 200, 500\mu$A. The comparative plots show quite a surprising degree of agreement, as seen in Figure 22, though again there is a tendency for the calculated values to be too high in frequency as the current I_f gets bigger.

Now we come to the anomaly hinted at above: if we use a diode at the top of the ladder which is *different* from the transistor type used for the input pair, then we get *more gain* than is predicted by the model. This is shown in Figure 23, where the top pair of diodes have been replaced with 1N4148 types ($I_f = 100\mu$A and no resonance, i.e. $k = 0$): the difference between the curves in the passband is about 6dB; the curve at approximately 0dB, which uses all the same transistors/diode-connected transistors is what the model predicts (as shown by the blue trace). All manner of different behaviours can be seen, curves shifting up or down, by replacing different pairs of transistors in the ladder for 1N4148 types. I have not had the opportunity to get fully to the bottom of this, but my suspicions are that the approximations using the standard 'cut-down' Ebers-Moll equations, at equations (1) and in particular (26), are too simplistic: I

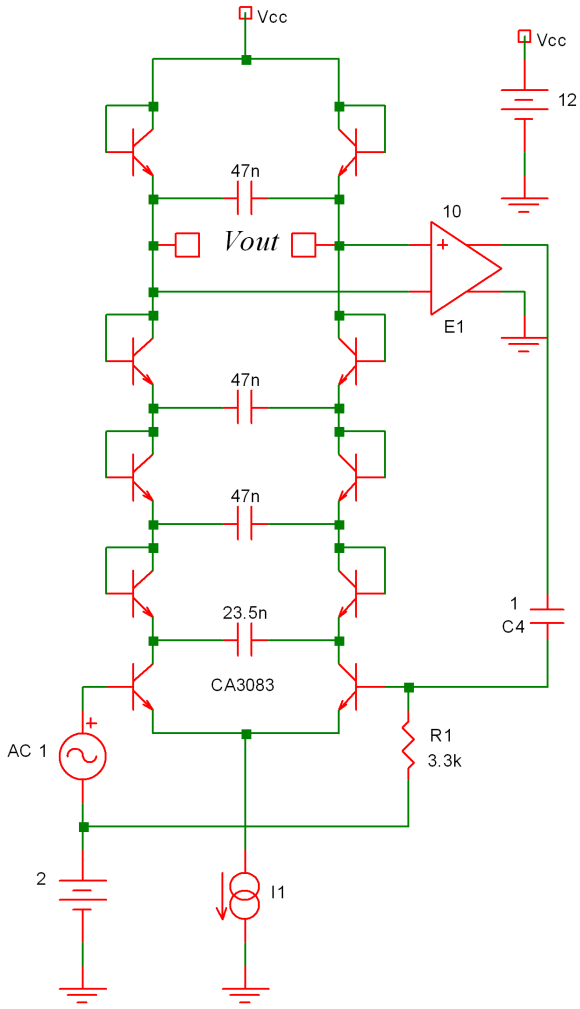


Figure 21: Simulation circuit for TB-303-type filter accuracy check

think maybe other parameters like the 'emission coefficient' may need to be included, so that when the same transistors are used these cancel, but when we mix diodes and transistors, they do not, and we get an extra gain factor, similar to the '3' derived at equation (28) (as the '3', it would appear in the denominator of the exponential power, and would generally be greater than 1).

As a second check of a simulation against the models, use an EMS-type filter: add another pair of diodes in at the top of the ladder (in Figure 21, so again all diode-connected CA3083, so as to avoid the previous problem); change all *four* capacitor values to 15nF (the frequency response comes right down otherwise); set no resonance, $k = 0$; and use the same values for I_f. The expression used for the calculations is:

$$20 \log \left| H_{ems} \left(\frac{3^{\frac{1}{4}} 4 C V_T}{I_f} 2\pi f j, k \right) \right|,$$

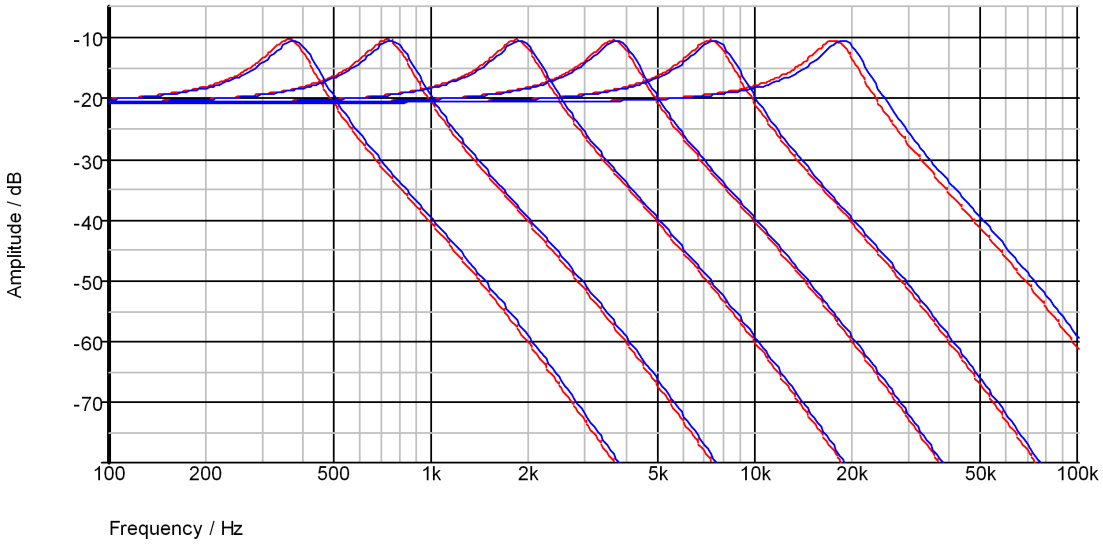


Figure 22: Plots from simulation of TB-303-type filter, red, vs. calculation of H_{tb}, blue

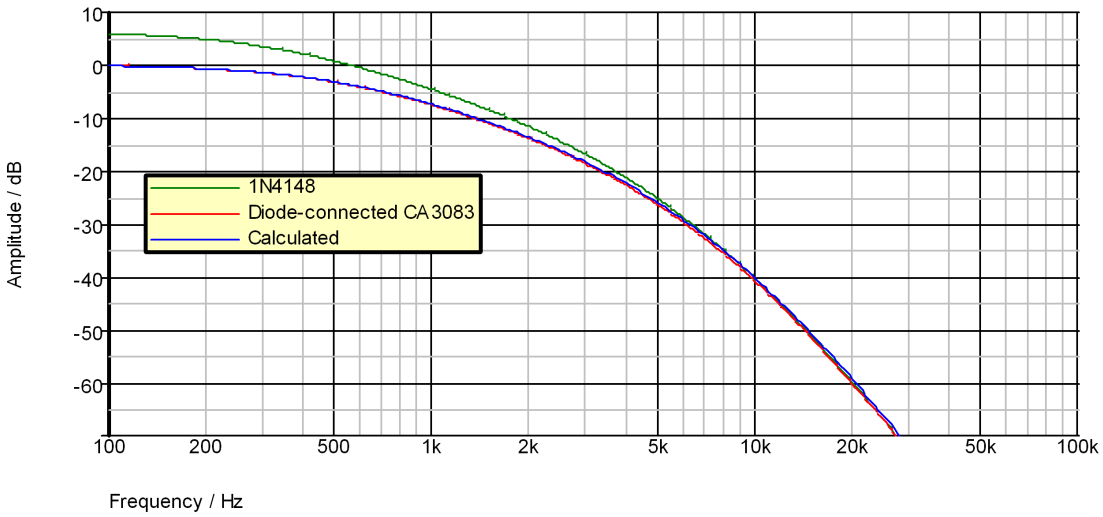


Figure 23: Effect of diode vs. diode-connected transistor at top of ladder

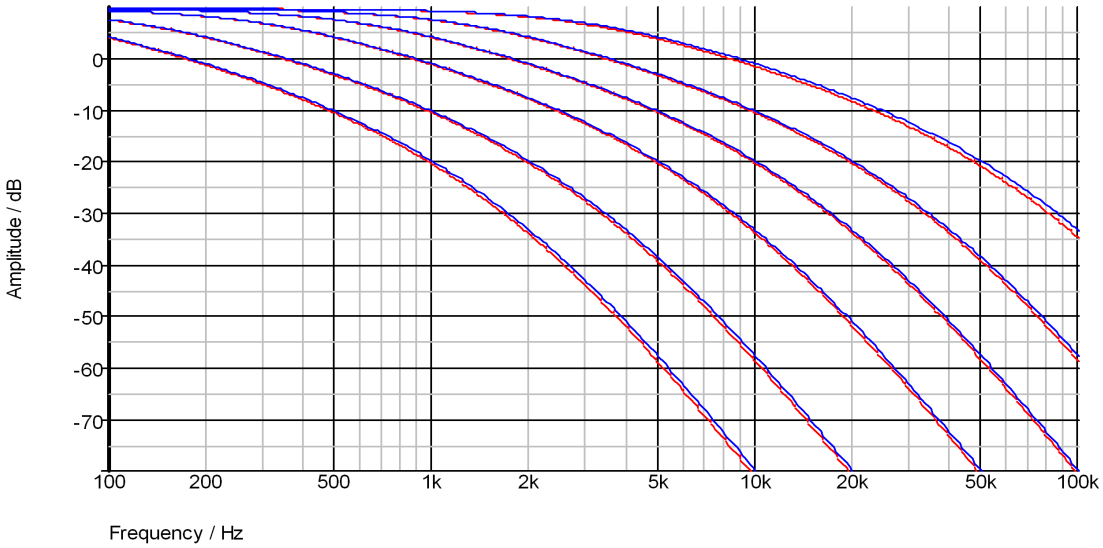


Figure 24: Plots from simulation of EMS-type filter, red, vs. calculation of H_{ems}, blue

and the results are shown in Figure 24: again they show good agreement, but the point to note is that the passband now shows the extra 9.5dB $\equiv \times 3$ gain as predicted by the model, and caused by the (now) three pairs of diodes at the top of the ladder.

3.5 Poles and frequency responses

Of the four functions developed in Section 3.3, in this section we will concentrate mainly on the simplest, $H_1(s)$: whilst there are differences between the four, sometimes they are not great, and in general they all share the same characteristics.

First we look at the poles of $H_1(s)$: at $k = 0$, these all lie on the real axis, paired as $(-3.532, 0)$ and $(-2.347, 0)$, and $(-1.000, 0)$ and $(-0.121, 0)$. As k increases, both pairs move toward each other, until when k is about 1.07, the right-hand pair meet at $(-0.48, 0)$, and then they become complex and break away from the real axis. At $k \approx 2.04$, the left-hand pair meet at $(-3.076, 0)$, and thereafter they too become complex and break away from the real axis. As k increases the poles arch away from each other: at $k \approx 18.4$, the right-most pair reach the imaginary axis, at which point the filter will be oscillating—this pattern is shown in Figure 25[†]. The following Figure 26 shows these

[†]As with H_{std}, a similar trick was used to get a nice-looking spacing between the poles for the plot—k was varied as $18(n/30)^4$ for $n = 0, 1, \ldots 30$, so the actual values for k are: 0.0, 0.0000222222, 0.000355556, 0.0018, 0.00568889, 0.0138889, 0.0288, 0.0533556, 0.0910222, 0.1458, 0.222222, 0.325356, 0.4608, 0.634689, 0.853689, 1.125, 1.45636, 1.85602, 2.3328, 2.89602, 3.55556, 4.3218, 5.20569, 6.21869, 7.3728, 8.68056, 10.155, 11.8098, 13.659, 15.7174, 18.0

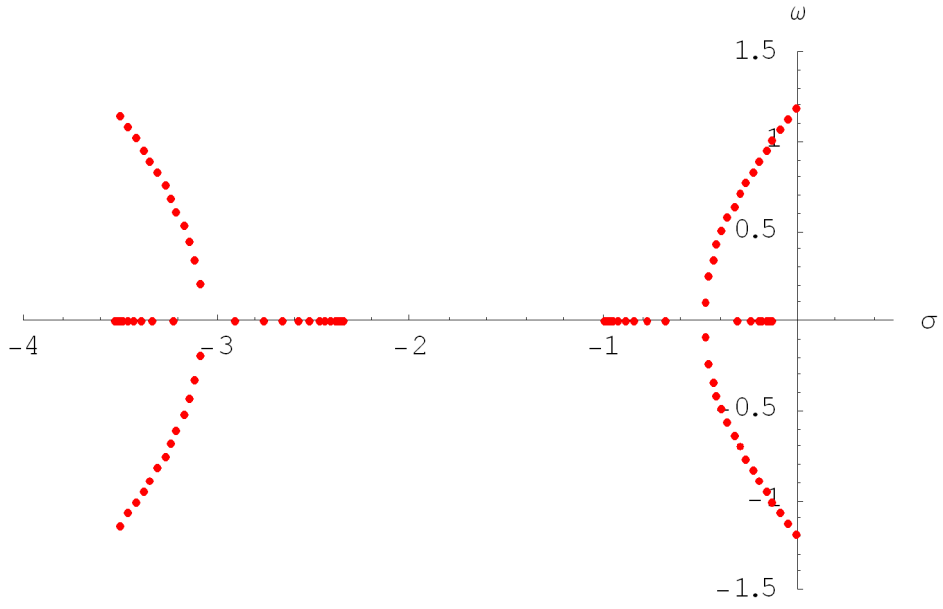


Figure 25: Poles of the transfer function $H_1(s)$

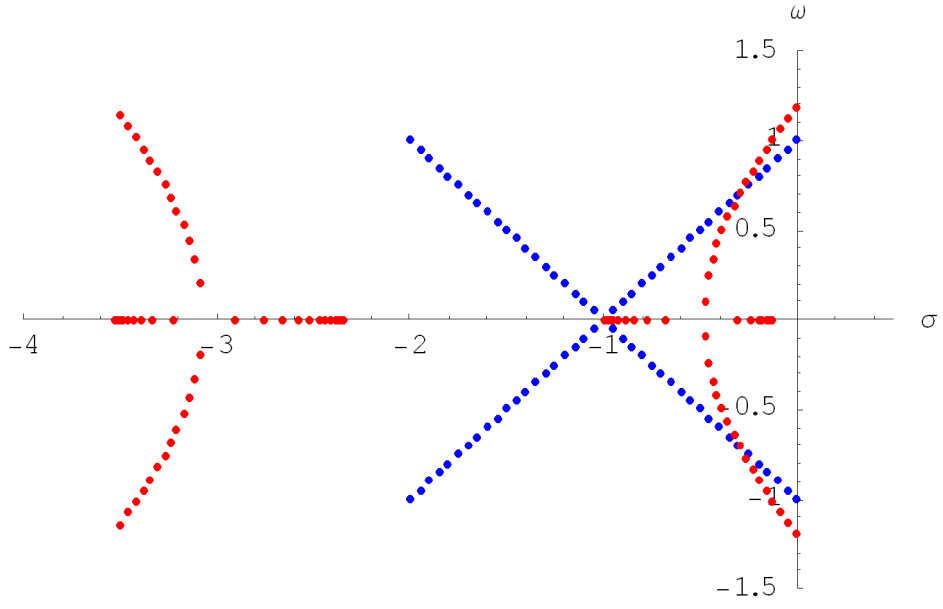


Figure 26: Poles of $H_1(s)$ compared to those of $H_{std}(s)$

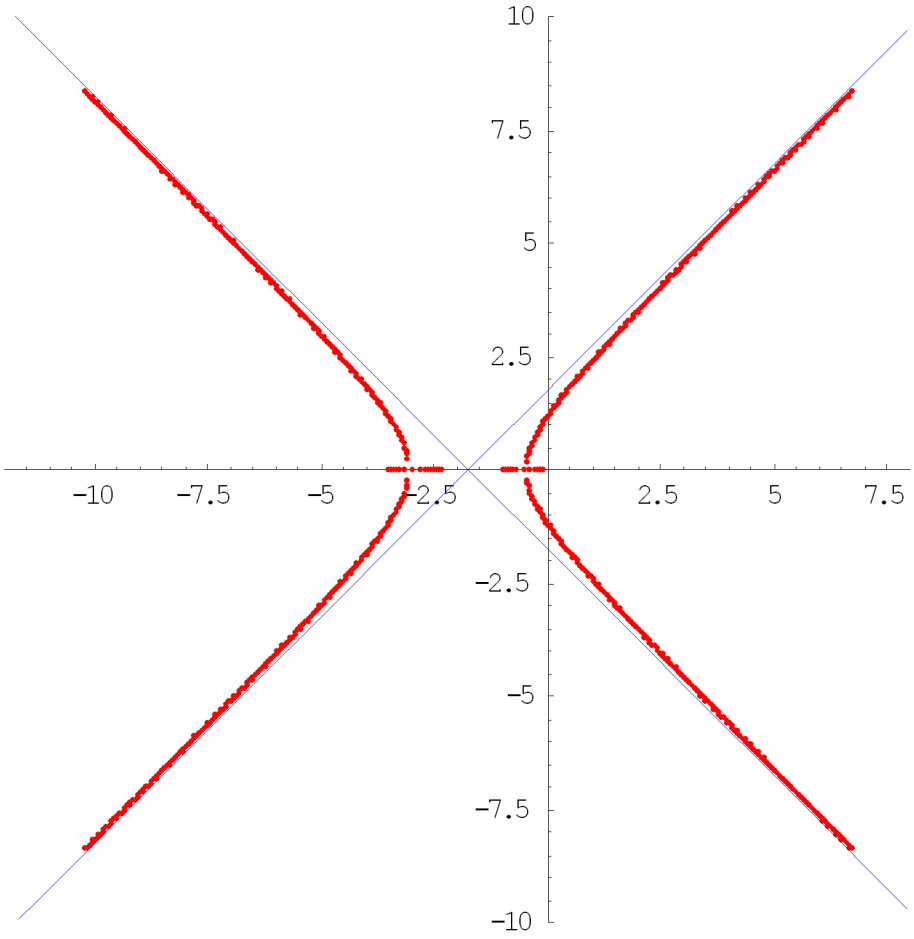


Figure 27: Asymptotes to poles of $H_1(s)$ for large k

poles overlaid with those of the Moog transistor ladder, from function H_{std}, which shows quite a difference: the initial nearness of the right-most diode pole to the imaginary axis (at $k = 0$) means that the passband drops off much sooner than the Moog filter, and when combined with the increased distance to the left-most pair means that the 'corner' is much less pronounced; also note that the point where the diode poles cross the imaginary axis is a little offset above $\omega = 1$—this is reflected in later curves of the amplitude response, where the peaks are offset, and also the point where the abrupt change in phase occurs in the phase response.

Whilst looking at some general feedback and control systems theory I discovered that the 45° 'X' pattern of the H_{std} poles isn't just a facet of this (fairly standard) transfer function, but it is actually a facet of four-pole filters *per se* (and specifically four poles and no zeroes). This can be seen if we see what happens in the diode case when the poles have moved much further away from the real axis: at large distances they become asymptotic to a pair of lines that indeed cross in an 'X', as is shown in Figure 27, which plots the poles out to about $k = 20,000$. The point where the asymptotes

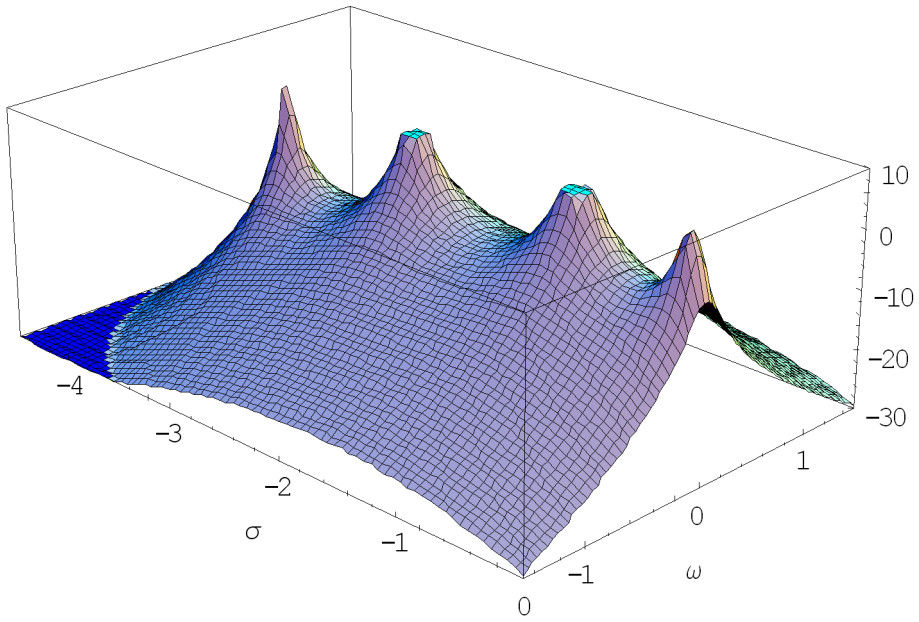


Figure 28: The surface $20\log|H_1(s)|$ at $k = 0$

cross is known as the 'centre of asymptotes', and is given (in this case) by the mean of the poles at $k = 0$, i.e.

$$\sigma_c = \frac{-3.532 - 2.347 - 1.000 - 0.121}{4} = -1.75.$$

Looking at the surface $20\log|H_1(s)|$ for $k = 0$ in Figure 28 shows how the poles are strung out along the real axis, and the nearness of that at $(-0.121, 0)$ to the imaginary axis, makes the frequency response (the slice down the imaginary axis) fall off much sooner than that for the transistor ladder (though as before, note the scaling along the imaginary axis is linear). The next Figure 29 shows the surface at $k = 1.6$, so the right pair of poles has just broken away from the real axis, but the left pair are still on it: this is starting to look more similar to the transistor one, and note that the passband has also started to droop well below 0dB, as the transistor one does too.[†]

The amplitude response with frequency is shown in Figure 30: this shows the same general trends as $H_{std}(s)$ of the Moog ladder, but note the greater variation over the vertical scale (here I have plotted to -50dB; the equivalent Moog one is only to -40dB, and appears considerably less 'spread out' vertically). (The values of k are $18n/4$, where $n = 0, 0.5, 1, 1.5, 2, 2.5, 3, 4$, and so are sort of proportionally the same as the Moog plot.) As mentioned previously, also note that the peaks tend to come to a point at a value of ω just above 1. The quicker entry into the stopband, and the more gentle roll-off as we

[†]Again, some animated gifs of these surfaces may be found at my website at: `http://www.timstinchcombe.co.uk/synth/poles/poles.html`

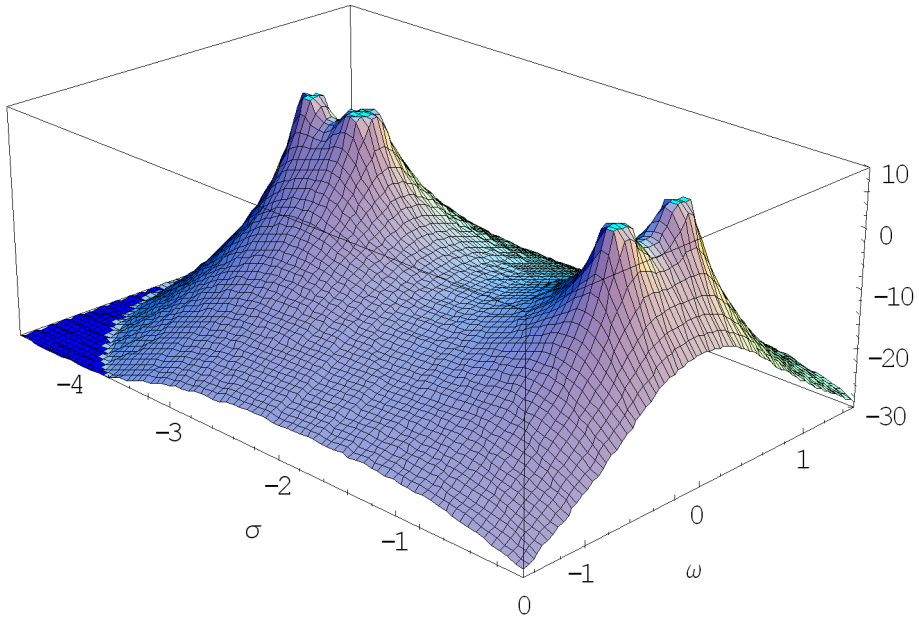


Figure 29: The surface $20\log|H_1(s)|$ at $k = 1.6$

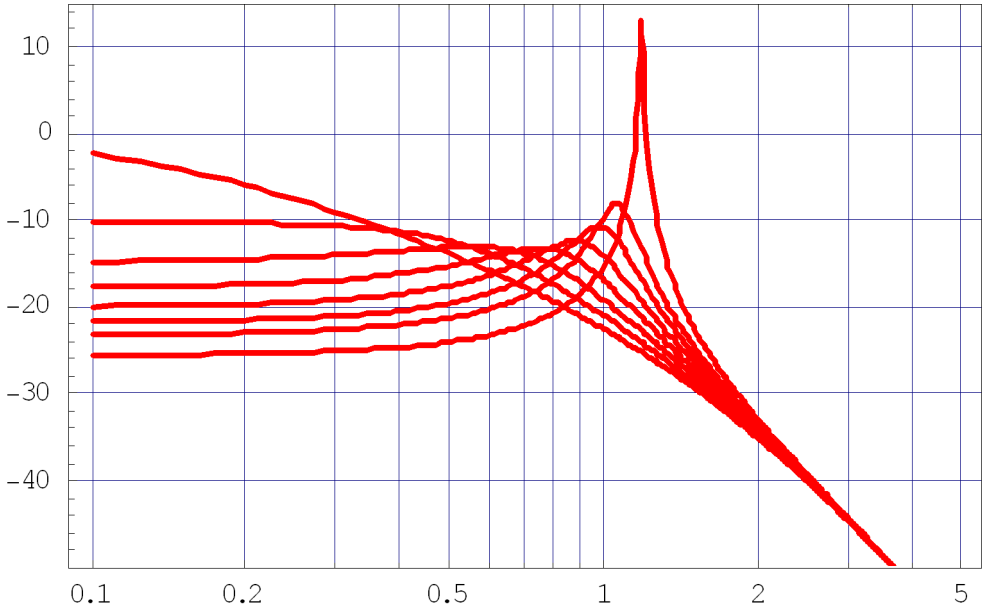


Figure 30: Magnitude responses of $H_1(s)$, various k

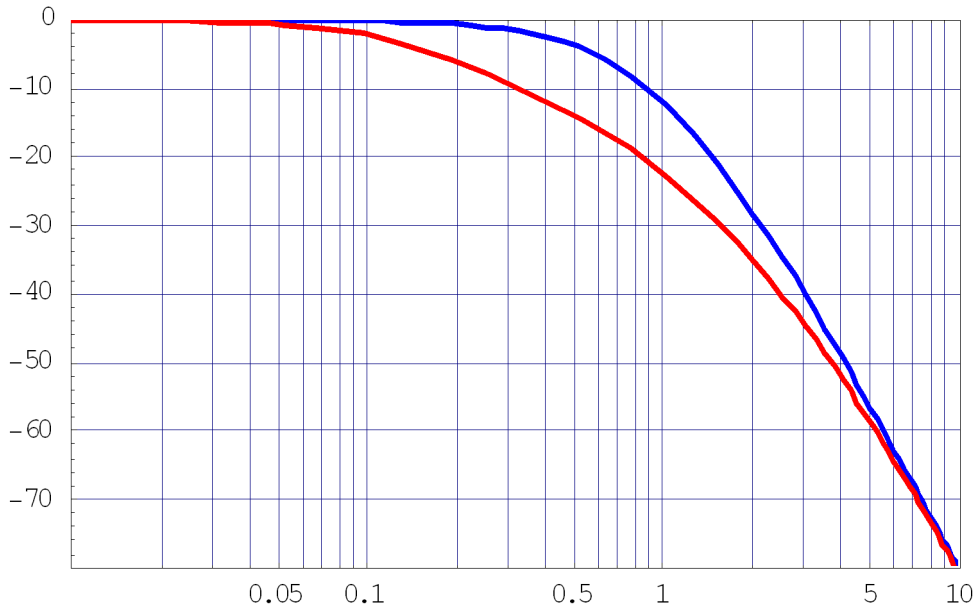


Figure 31: Comparison of diode (red) and transistor (blue) magnitude responses

go 'round the corner' is emphasized in Figure 31, which shows the no-resonance, $k = 0$, plots for both the diode ladder and transistor ladder together: it is this gentler slope that causes the diode ladder to sometimes be referred to as an '18dB filter'—a gentler slope it may be, but calling it a '3-pole, 18dB/octave' is at best misleading, and at worse simply incorrect; further into the stopband it clearly re-joins the transistor curve, showing its proper 4-pole, 24dB/octave descendancy.

Phase responses of $H_1(s)$ are shown in Figure 32, and are broadly similar to the Moog equivalent, apart from the shift in the position of the abrupt phase change at high k, as already noted.

We now move on to consider some of the differences between the various diode ladder configurations, starting with the effect of the three diodes at the top of the ladder (EMS-type) versus just the one ('standard' and TB-303-type). Because of the way that the $d = 3$ factor affects the denominator of the transfer function expression (30) in forming $H_{ems}(s)$, there is a significant impact on the left-most, non-dominant pole pair, as can be seen when we overlay the poles of $H_1(s)$ and $H_{ems}(s)$ in Figure 33. The actual pole locations for $k = 0$ are:

$$\begin{aligned} H_{ems}: &\quad (-4.530, 0), (-2.756, 0), (-0.966, 0), (-0.083, 0) \\ H_1: &\quad (-3.532, 0), (-2.347, 0), (-1.000, 0), (-0.121, 0). \end{aligned}$$

The increased distance of the non-dominant pole pair from the imaginary axis means that the magnitude response of the EMS-type filter is quite a bit flatter, as is seen when this is compared against the standard case, shown in Figure 34. Note that to make this

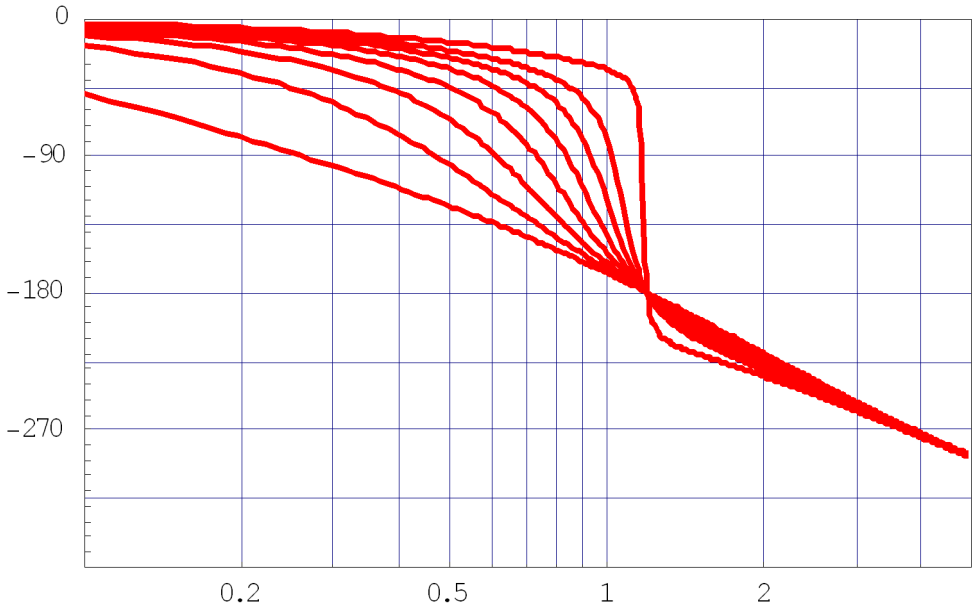


Figure 32: Phase responses of $H_1(s)$, various k

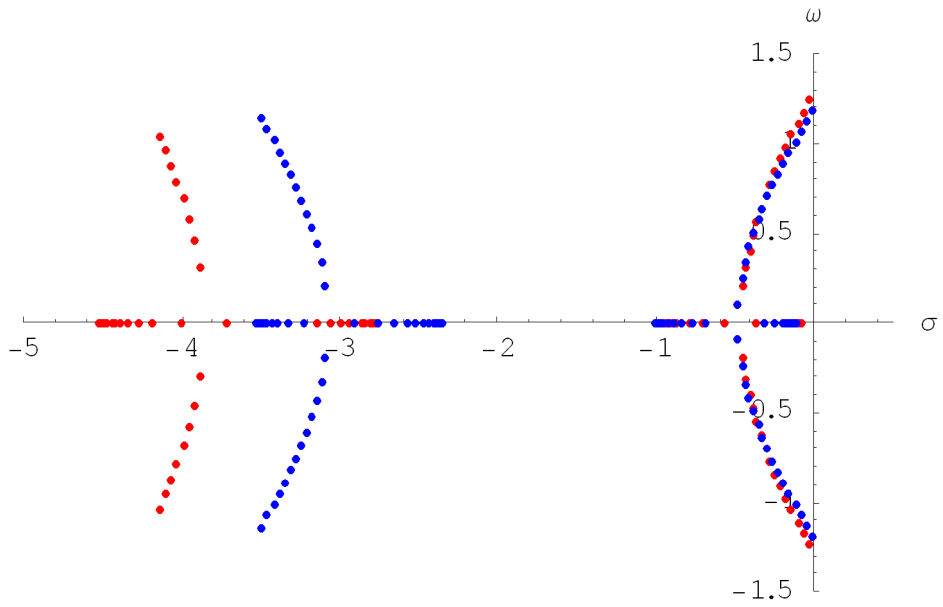


Figure 33: Poles of $H_{ems}(s)$, red, compared to $H_1(s)$, blue

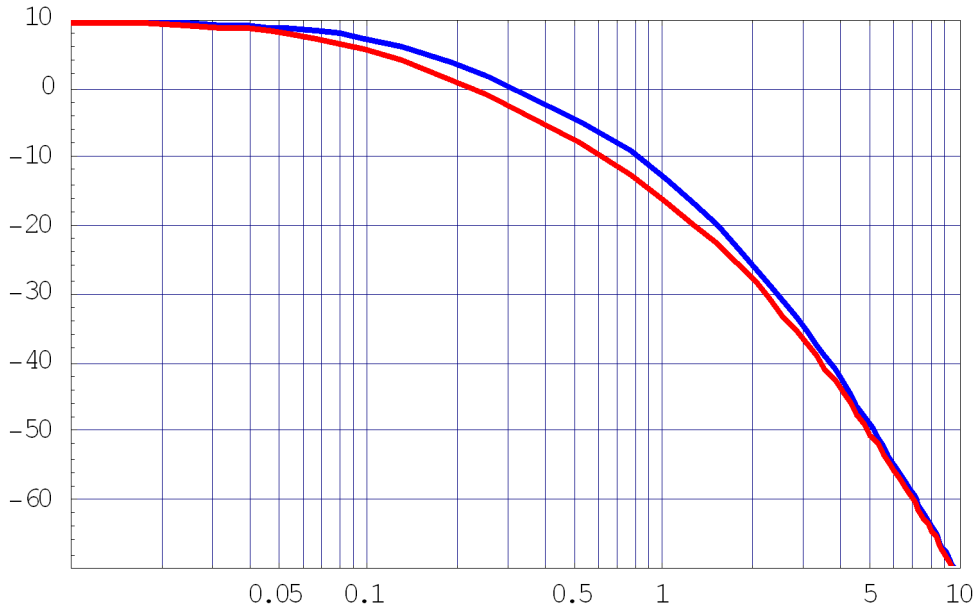


Figure 34: Magnitude responses: H_{ems} red, $H_1(s)$ (shifted), blue

a fair comparison, as the EMS-type filter has the extra gain due to the three diodes, the H_1 curve has had an extra 9.54dB ($\equiv \times 3$) added to it, to shift it up to be at the same point at the left of the plot. This change thus looks like a significant difference that might be worth having (in order to keep the filter 'different' from the competition).

Another minor difference for the EMS-type filter is that the feedback gain needed to make the filter oscillate is quite a lot less: the right-most poles reach the imaginary axis at a value of k of about 9.8, as compared to the value of 18.4 for H_1.

We now consider the effects of halving the value of the bottom capacitor in the ladder, by comparing the functions $H_{ems}(s)$ and $H_d(s)$, the latter having the configuration with the bottom capacitor halved. Intuitively, halving the bottom capacitor should increase the cut-off frequency: the impedance of the capacitor doubles, and so more of the signal (at a fixed frequency) will make it up the ladder to the output, rather than being shorted through the capacitor. The analysis bears this out: from expressions (32) and (33) for the cut-off frequencies of the two functions, at any particular I_f, we clearly have

$$\omega_{c_d} = \frac{I_f}{2^{\frac{7}{4}}3^{\frac{1}{4}}CV_T} > \frac{I_f}{3^{\frac{1}{4}}4CV_T} = \omega_{c_ems}.$$

However the normalization process will remove this shift, and looking at the amplitude responses at zero resonance, there is hardly any difference to be seen at all, as Figure 35 shows. This suggests that halving the capacitor causes a shift in the frequency response which might easily be removed by a trim pot for example, placing you back where you started from. I've seen the notion expressed in several places on the Internet

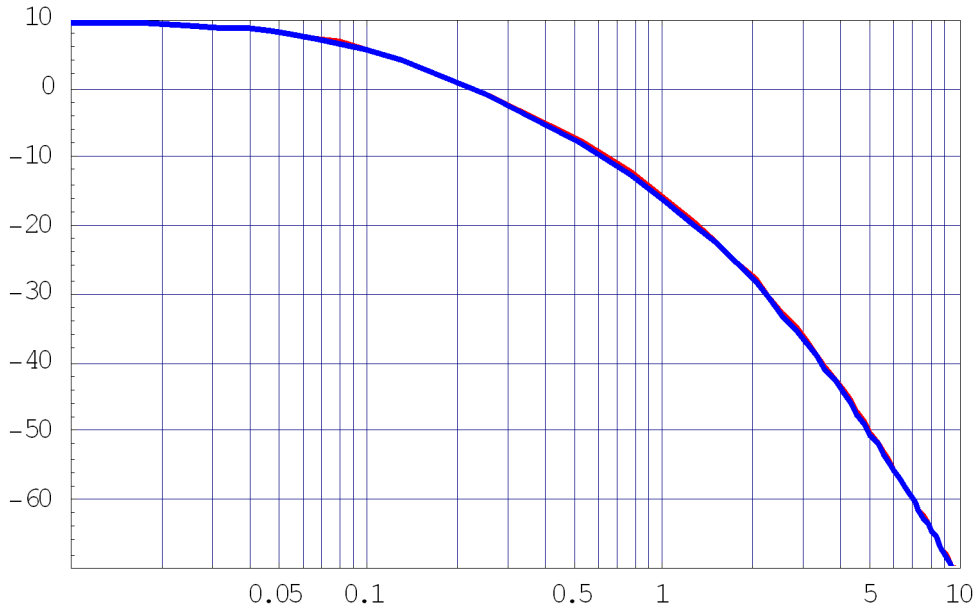


Figure 35: Comparison of magnitude responses: $H_d(s)$, red, and H_{ems}, blue

that halving the capacitor 'shifts it's pole up an octave', hence increasing the cut-off frequency. Simply examining the poles exposes the fallacious nature of such a claim, and again demonstrates why there is so little difference between the responses:

$$\begin{aligned} H_{ems}: &\quad (-4.530,0),(-2.756,0),(-0.966,0),(-0.083,0) \\ H_d: &\quad (-4.220,0),(-2.799,0),(-1.014,0),(-0.084,0), \end{aligned}$$

being for $k=0$—these and others for various k are shown in Figure 36. From this we see that all the poles have moved, the left-most pair more so than the right-most, which hardly move at all. If we take the transfer function of a fourth-order filter, double one of its poles, form the corresponding function and re-normalize it, the doubled pole ends up as $2^{\frac{3}{4}}$ times the original, and the others as $2^{-\frac{1}{4}}$ times their original values[†]. If we do this to each pole in turn of $H_{ems}(s)$, the values we get are:

$$\begin{aligned} &(-7.619,0),(-2.318,0),(-0.812,0),(-0.070,0), \\ &(-3.810,0),(-4.636,0),(-0.812,0),(-0.070,0), \\ &(-3.810,0),(-2.318,0),(-1.625,0),(-0.070,0), \\ &(-3.810,0),(-2.318,0),(-0.812,0),(-0.140,0). \end{aligned}$$

[†]If the poles are originally $-p_1,-p_2,-p_3$ and $-p_4$, and we double p_1 so that the denominator is now $(s/(2p_1)+1)(s/p_2+1)(s/p_3+1)(s/p_4+1)$, to re-normalize put $s'^4 = s^4/(2p_1p_2p_3p_4) = s^4/2$ as $\prod p_i = 1$ by the original normalization, so substitute $s = 2^{\frac{1}{4}}s'$, then drop the prime to get$(2^{\frac{1}{4}}s/(2p_1)+1)(2^{\frac{1}{4}}s/p_2+1)(2^{\frac{1}{4}}s/p_3+1)(2^{\frac{1}{4}}s/p_4+1)$—the factors giving the new poles are clear, and their product is again 1.

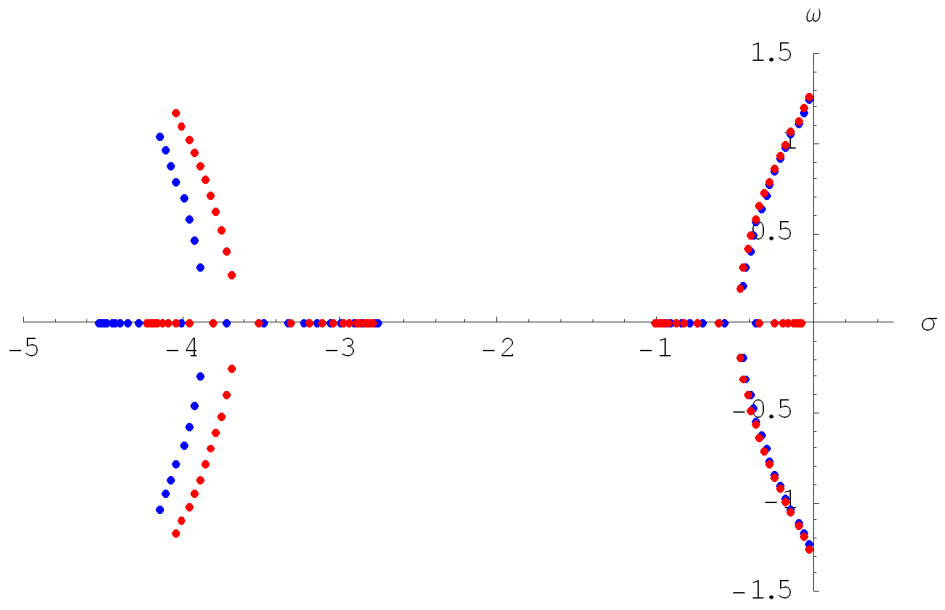


Figure 36: Poles of $H_d(s)$, red, compared to those of $H_{ems}(s)$, blue

It is clear that the poles of $H_d(s)$ look nothing like any of these patterns, thus apparently nullifying the notion at a stroke. The reason is also fairly obvious: unlike the Moog transistor ladder, where the filter sections are effectively buffered from each other, in a diode ladder this buffering effect is just not there, and hence the poles do not exhibit the kind of independence that would be necessary to support the notion.

The question then remains as to why some designers *have* decided to halve the value of the bottom capacitor. At the moment I can only speculate on such a motive, but I suspect that it may be due to some sort of stability issue of which I am currently unaware, and hence that I'm certainly not in a position to comprehend!

3.6 Later EMS 5-pole filter with zero

It was mentioned earlier that the later EMS filter is actually a five-pole design, and that the resistor in series with the extra capacitor gives the transfer function a zero. We will have a quick (read 'non-rigorous'!) look at how this arises. First, simple simulation quickly reveals the presence and rough location of the zero, as seen in Figure 37 (at $I_f = 50\mu$A and no resonance). It can be seen that the 5-pole trace (green), turns back up at around the 10kHz mark, just as it is starting to distinguish itself from its 4-pole brother (the red trace), and at higher frequencies, the two lines appear to be parallel. It thus looks like the 100Ω resistor combines with the extra 100nF capacitor to form a zero: indeed, calculating

$$f = \frac{1}{2\pi RC}$$

with these values yields 15.9kHz, which seems to back up the assertion.

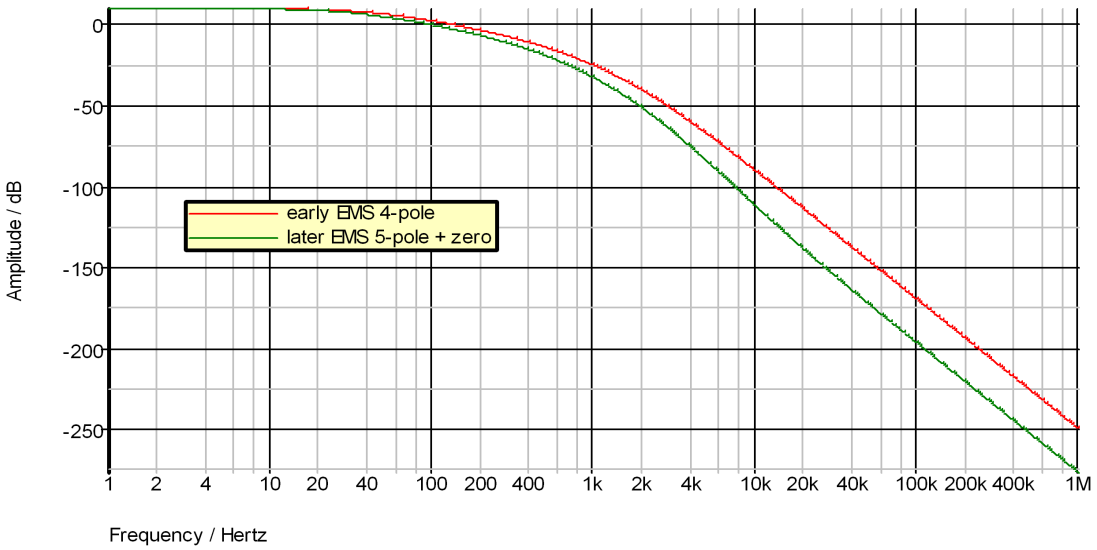


Figure 37: Early EMS 4-pole response vs. later 5-pole with zero

To get a *very rough* feel for how this happens, I looked at what the transfer function for a single *transistor* stage might look like with a resistor in it, by substituting for I from

$$\Delta V_n = I\left(R + \frac{1}{sC}\right)$$

rather than

$$\Delta V_n = \frac{I}{sC}$$

as used in equation (10), to get

$$\Delta I_{n+1} = \Delta I_n + 2\frac{sC}{sCR+1}\Delta V_n$$

instead. Then when ΔV_n is eliminated and the terms re-arranged, rather than the '$1/(s+1)$' of equation (11) we get

$$\frac{\Delta I_{n+1}}{\Delta I_n} = \frac{sCR+1}{sC\left(\frac{4V_T}{I_f}+R\right)+1},$$

which shows a zero at the frequency we guessed above, and clearly through all the machinations in arriving at the transfer function (17), assuming it is at stage 1 (or in fact any stage), this zero will remain in the numerator, and so be a zero for the whole thing. I suspect the mechanism for introducing the zero in the *diode* ladder is extremely similar, despite all the lack of buffering etc., and so I am fairly certain that something similar to this is how it arises there too.

4 Conclusions

We have analysed the Moog transistor ladder filter and several of the derivative diode ladder filters, derived their transfer functions, and looked at their main characteristics, such as poles and frequency responses etc. Whilst this is interesting for its own sake, it doesn't necessarily tell us how to design filters that are musically interesting: gross differences in these characteristics are likely to indicate that the filters will sound different, but similarities may not be an indicator that two filters will sound the same. For example I suspect the ancillary envelope circuitry around the TB-303 filter, with all its complex interactions, plays a major role in shaping what people regard as the 'legendary sound' it makes, as much or maybe more than, the actual filter core examined here.

The diode ladder model suffers from some issues with regards to the matching (or otherwise) of the parameters of the diodes in the ladder versus the transistors in the input and/or output stages, but my suspicions are that these could probably be accounted for easily enough. The true motive for the sometimes-seen halving of the bottom capacitor (in the diode ladders particularly) is still not clear, as from what has been presented here, it seems to have little impact on the filter responses.

References

[1] R.A. Moog, *A Voltage-Controlled Low-Pass High-Pass Filter for Audio Signal Processing*, Audio Eng. Soc. preprint 413, 17th AES Annual Meeting, 11-15 Oct, 1965

[2] R.A. Moog, *Electronic High-pass and Low-pass Filters Employing the Base to Emitter Diode Resistance of Bipolar Transistors*, USPTO No. 3,475,623, Oct 28, 1969

[3] T.E. Stinchcombe, *Derivation of the Transfer Function of the Moog Transistor Ladder Filter*, 15 Feb 2004 (updated 5 July 2005). (Previously at `www.timstinchcombe.co.uk`.)

[4] T. Stilson and J. Smith, *Analyzing the Moog VCF with Considerations for Digital Implementation*, International Computer Music Conference, Hong Kong, 1996, (`http://ccrma-www.stanford.edu/~stilti/papers/moogvcf.pdf`)

[5] A. Huovilainen, *Non-Linear Digital Implementation of the Moog Ladder Filter*, 7th Int. Conf. on Digital Audio Effects (DAFx'04), Oct 5-8, 2004, Naples, Italy (`http://dafx04.na.infn.it/WebProc/Proc/P_061.pdf`)

[6] T. Helie, *On the use of Volterra Series for Real-Time Simulations of Weakly Nonlinear Analog Audio Devices: Application to the Moog Ladder Filter*, 9th Int. Conf. on Digital Audio Effects (DAFx'06), Sept 18-20, 2006, Montreal, Canada (`http://dafx.ca/proceedings/papers/p_007.pdf`)

[7] M. Civolani and F. Fontana, *A Nonlinear Digital Model of the EMS VCS3 Voltage-Controlled Filter*, Proc. of the 11th Int. Conference on Digital Audio Effects (DAFx'08), Espoo, Finland, September 1-4, 2008 (`http://www.acoustics.hut.fi/dafx08/papers/dafx08_06.pdf`)

[8] F. Fontana, *Preserving the Structure of the Moog VCF in the Digital Domain*, Proc. Int. Computer Music Conference, Copenhagen, Denmark, August 27-31, 2007

[9] T.H. Wilmshurst, *Analog Circuit Techniques with Digital Interfacing*, Newnes, 2001

[10] A.S. Sedra and K.C. Smith, *Microelectronic Circuits*, Oxford University Press, 4th edn., 1998

[11] M.K. Berry, *Electronic Music and Creative Tape Recording*, BP51, Bernard Babani (Publishing) Ltd., 1978