Reducing low frequency non-linearities in loudspeakers with signal processing

Introduction

Almost all kinds of loudspeakers and any other form of actuators which produce sound or vibrations behave differently dependant on their amplitudes, most notably in moving coil loudspeakers. This is most notable at low frequencies which shows up in harmonic and intermodulation distortions. The main reasons for these nonlinearities are the following:

- The force factor (Bl(x)) being applied to the moving voice coil changing with the excursion
- The electric self-inductance also depends on the voice coil excursion (L(x))
- The non-linear suspension stiffness applying a restorative force with excursion (k(x))

One way in which these problems could be addressed is to attempt to carefully design each of these components to limit their non-linear affects, however this would likely have a very high associated cost to the development of the solutions and would likely be impossible to perfect. A different approach is to predistort the input signal with an appropriate nonlinear filter such that it significantly reduces the distortions caused by these nonlinearities.

Background

Causes of Distortion

Force Factor:

The force factor (Bl(X)) describes the coupling between the mechanical and electrical side of an electromagnetic transducer. It can be described as the integral of the flux density versus the coil length cutting the flux, l. Clearly as the windings leave the gap the force factor decreases; therefore, the force factor is not a constant, but instead changes with the voice coil displacement, x, in a non-linear fashion. This can be clearly shown from figure 1 which shows typical force factors at displacement for both an overhang motor configuration, where h_{coil} is greater than h_{gap} as shown in figure 2, and where h_{coil} is an equal length to h_{gap} (Klippel 2005a). This non-linearity is frequency independent and solely dependent on the displacement of the voice coil and the motor geometry (Klippel 2005a).

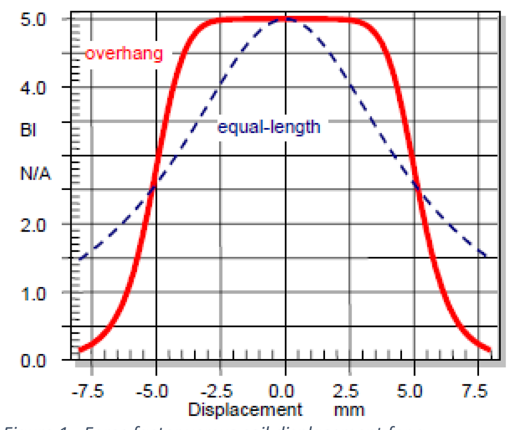


Figure 1 - Force factor versus coil displacement for a loudspeaker with an overhang motor arrangement and an equal length arrangement (Klippel 2005a)

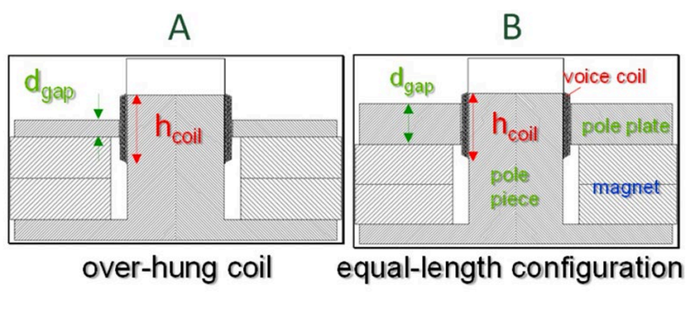


Figure 2 - Typical motor arrangements with an overhang configuration and an equal length configuration (Klippel 2017)

Electrical Self-Inductance:

The inductance of the voice coil of a moving coil loudspeaker is in part determined by the permeability of the surroundings. For this reason, as the displacement of the voice coil changes, so does the inductance, as when the voice coil has negative displacement (has gone towards the base of the magnet) the coil is surrounded by the more permeable magnet and iron surrounds so therefore has a higher inductance as a result of the lower magnetic resistance. This can be seen from figure 3 which shows the input impedance of a moving coil loudspeaker which is free standing and clamped at +/- 7mm (Klippel 2005a).

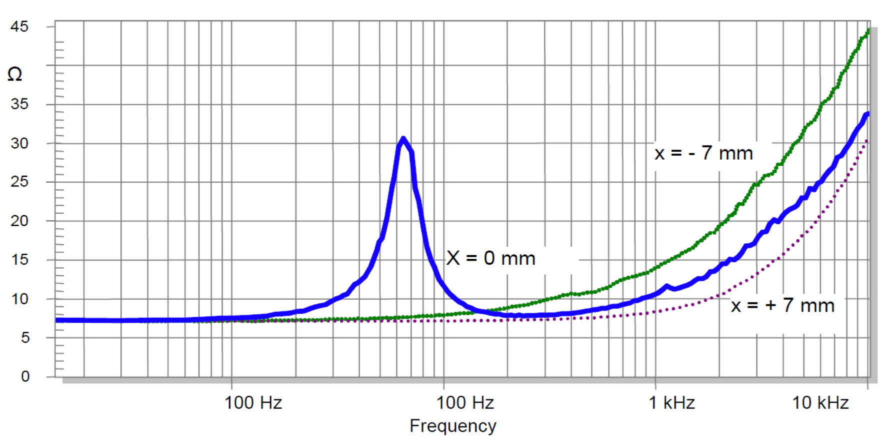


Figure 3 - Input electrical impedance of a moving coil loudspeaker free floating and clamped at +/- 7mm (Klippel 2005a)

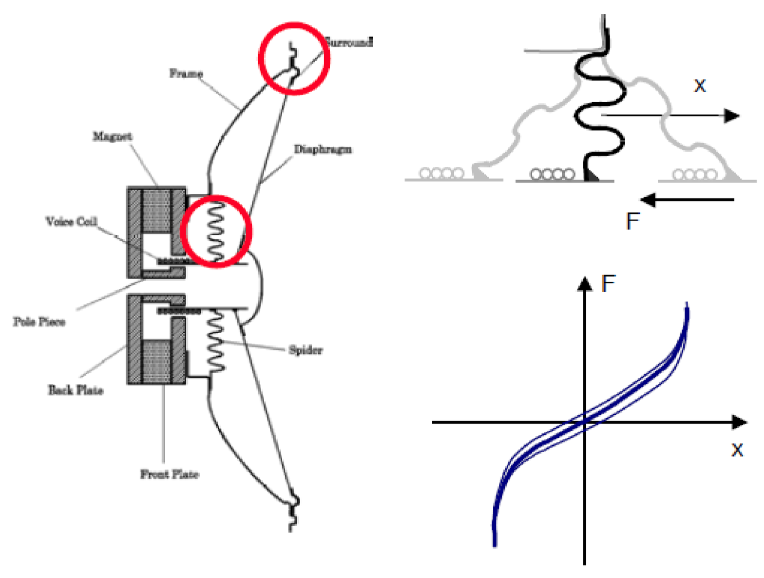


Figure 5 - Suspension system in a typical moving coil loudspeaker and force reflection curve (Klippel 2005a)

The clamped measurements don't have the mechanical resonance that can be seen on the free-standing measurement, which is due to the voice coil being restrained in a fixed position (Klippel 2005a). As can be seen by looking at the high frequency increase in impedance, the inductance can be seen to clearly increase as the negative displacement increases.

As well as the inductance being dependant on the displacement of the coil, the inductance is also dependant on the input current. When the current has a high AC frequency the impedance is increased as can be seen from figure 3, however it does not increase in a linear manner. This is due to the loses in the eddy currents known as para-inductance, which can be modelled sufficiently accurately as a second inductor, which has been shunted by a resistor, in series with the first inductor (Dodd *et al.* 2004). This can be seen from the electrical equivalence circuit in figure 4 as L_2 and R_2.

Non-linear Suspension Stiffness

Most moving coil transducers use a suspension system to centre the voice coil in the gap and apply a restoring force which moves the coil back to the rest position. This is usually comprised of a spider attaching to the voice coil former and the surround which is attached to the outer edge of the cone. These are typically made of an impregnated fabric, rubber or plastic moulded to a particular shape. This suspension behaves like a normal spring and can be represented by a force reflection curve such as shown by figure 5.

As can be seen from the force reflection curve in figure 5, the restoring force from the surround is only linear at a relatively low displacement. Towards the maximal displacement the surround increases the resistance the more strained it becomes meaning that unless more force is applied the displacement is less than that a linear model would suggest. When an AC force is applied, such as that produced by the motor in a moving coil loudspeaker, the displacement has a hysteresis caused because of the internal losses in the surround materials, which furthers the non-linear behaviour of the surround (Klippel 2005a).

Previous Approaches

Volterra Series

One approach that has been taken to reduce non-linearities has been to model the non-linear behaviour of a loudspeaker using a Volterra series expansion. This approach assumes that a moving coil loudspeaker is a non-linear time invariant system, however requires both the non-linearities and the input signal to be sufficiently small so that the convergence of the Volterra series can be guaranteed. This approach, as first developed by Kaizer, treats the entire system as a whole with lumped parameters (Kaizer 1987). This approach was later developed by Walter frank and others to create a linearising filter based off of this Volterra series expansion technique to reduce non-linearities (Frank *et al.* 1992).

The goal of these filters is to linearise the input signal in such a manner that the linear part of the transfer function through the system (filter and loudspeaker) remains linear, whilst the non-linear parts are compensated in such a way in the filter so that they are linearised at the output of the system as a whole.

Frank and his team modelled the moving coil loudspeaker as a non-linear, time invariant and causal system. They also assumed that the Volterra kernels were symmetrical, meaning that the permutations of the arguments were not important. They also concluded from measurements of harmonic distortion, as shown from figure 6, that only the first three harmonics have a significant impact on the loudspeaker non-linearities, therefore they used a third order Volterra filter which would cover the first three harmonics, whilst saving on superfluous processing. Rather than measuring the sound pressure, they used a laser vibrometer to measure the acceleration of the diaphragm of the loudspeaker under test (DUT) which is nearly proportional to the sound pressure, especially at low frequencies where these non-linearities are most prominent (Frank *et al.* 1992).

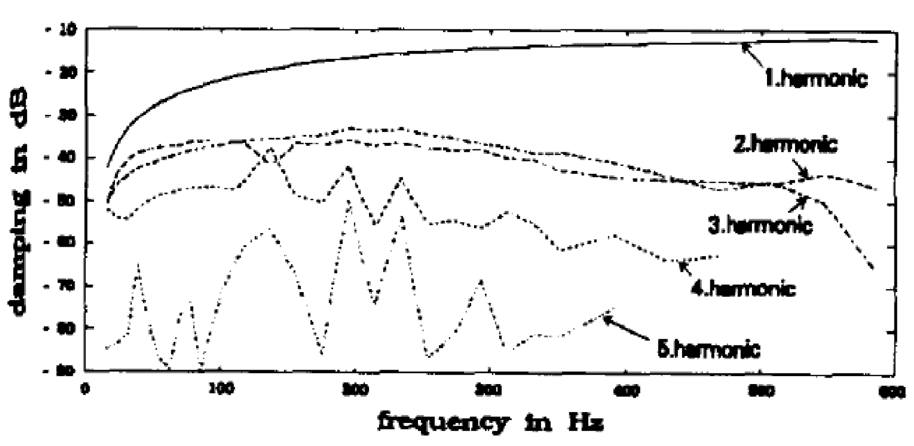


Figure 6 - Harmonic distortions of a moving coil loudspeaker

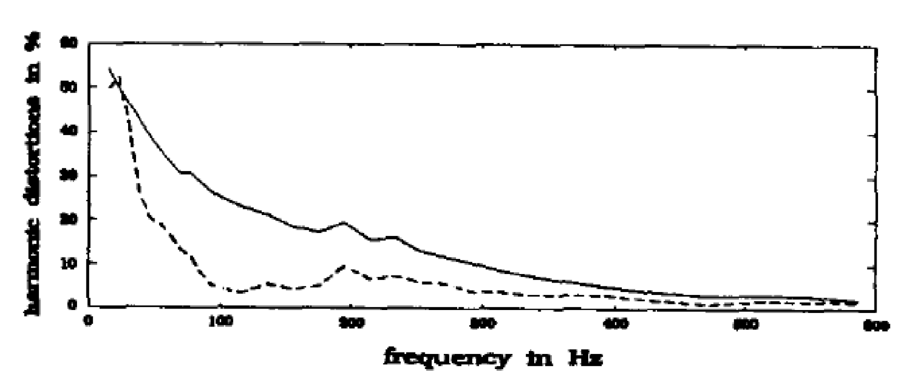


Figure 7 - Harmonic distortion in % without (solid line) and with (dashed line) linearisation from use of a Volterra series (Frank et al. 1992)

In order for such a filter to be useful, the kernels of the Volterra series being used as the basis for the filter must correspond to the loudspeaker. As such a way to estimate the Volterra kernels of the unknown system is required. This is known as system identification and will be covered more in depth in a later section, but Frank and his team explored two methods in order to achieve this. Their first proposed method was to identify using a zero mean gaussian noise and orthogonal functions, alternatively they proposed that a least mean square algorithm could be used in order to adaptively identify the correct Volterra series (Frank *et al.* 1992).

Once the system had been correctly identified they were able to apply an appropriate filter using the estimated kernel values meaning that the harmonic distortion was greatly reduced, specifically at the lower frequencies where this was targeted. This can be clearly seen by examining figure 7, which shows the results of their testing.

One of the downsides to this method is that by reducing the distortions in this manner you introduce more non-linearities at higher harmonics, however these tend to be negligible in comparison to the harmonic non-linearities that they reduce. This is due to the Volterra series having too few kernels to properly create an inverse non-linear filter. To fix this would require using more kernels which would increase both the identification process time and the latency of the filter. On top of this the Volterra series that it uses to perform the linearisation within the filter become unstable with high signal levels, which is where the causes of distortion we have looked at earlier become considerably worse, subsequently this method is less desirable for high signal level applications.

One of the downsides to this method is that by reducing the distortions in this manner you introduce more non-linearities at higher harmonics, however these tend to be negligible in comparison to the harmonic non-linearities that they reduce. This is due to the Volterra series having too few kernels to properly create an inverse non-linear filter. To fix this would require using more kernels which would increase both the identification process time and the latency of the filter. On top of this the Volterra series that it uses to perform the linearisation within the filter become unstable with high signal levels, which is where the causes of distortion we have looked at earlier become considerably worse, subsequently this method is less desirable for high signal level applications.

Mirror Filter

In contrast to the Volterra serries approach, where the system as a whole is lumped together in a closed loop system, Klippel's mirror filter instead looks to account for each non-linear component.

Because the behaviour of a moving coil loudspeaker at low frequencies can be modelled by an electromechanical equivalence circuit, you can model the effect of each component on the overall signal independently of one another. Being that the electrodynamic force factor (Bl(x)), the inductance of the voice coil (L(x)) and the stiffness of the driver's suspension (k(x)) are all dependant on the instantaneous displacement of the voice coil (x(t)) and as we have explored in previous sections, we therefore know these elements behave non-linearly and must therefore be modelled as such. In linear modelling this is often overlooked and as such are assumed to be a constant and are typically represented as Bl_0, L_0 and k_0, which is only valid for small displacements (Klippel 1992).

Klippel realised that the behaviour of a voltage driven electrodynamic moving coil loudspeaker could be represented by the following non-linear differential equation, as derived from the electrical equivalence circuit.

$$Bl(x) \cdot u_L(t) = R_e \cdot L^{-1}\{J(s)\} * x + R_e \cdot k(x) \cdot x + Bl(x) \cdot \frac{d[L(x) \cdot i_L(t)]}{dt} + Bl(x)^2 \cdot \frac{dx}{dt} - \frac{R_e}{2} \cdot i_L(t)^2 \cdot \frac{dL(x)}{dx}$$

Equation 1

Where R_e is the voice coil resistance, $i_L(t)$ is the input current, $u_L(t)$ is the voltage at the driver terminals and $J(s)$ is the mechanical and acoustic elements with constant parameters combined into an impedance which has been transformed into the Laplace domain. For a closed box this would be given as the following:

$$J(s) = m \cdot s_2 + R_m \cdot s + Z(s)$$

Equation 2

Where m is the moving mass, R_m is the mechanical damping and $Z(s)$ is the radiation impedance. As such the first term on the right-hand side is the only linear term and describes the effect of mechanical and acoustical elements with constant parameters. The impedance $J(s)$ is shifted into the time domain by the inverse Laplace transform and then subsequently convolved with the displacement. The second term represents the effects of varying stiffness of the suspension $(k(x))$, the third the impedance of the voice coil self-inductance, the fourth effect of electrical dampening on the force factor $Bl(x)$ and the final term, the electromagnetic force due to the varying inductance of the voice coil $L(x)$. These non-linear terms contain multiplications of time varying quantities which generate the harmonic and intermodulation distortions as a consequence of this.

By applying a non-linear filter to the input signal of the loudspeaker with an appropriate transfer function

$$u_L(t) = f[u(t)]$$

Equation 3

In front of the loudspeaker, the input signal to the driver $u_L(t)$ is predistorted so that the non-linear terms in equation 1 are compensated for. Therefore, the total system (the loudspeaker and filter) will have a behaviour that fits the following linear differential equation only using the constant parameters of the loudspeaker.

$$Bl_0 \cdot u(t) = L^{-1}\{R_e \cdot [J(s) + k_0] + sBl_0^2\} * x(t)$$

Equation 4

By inserting the transfer function into the non-linear equation, the difference between equation 1 and equation 4 can be made to yield the transfer function of the filter:

$$u_L(t) = \left\{u(t) + N_k(x) \cdot x + N_D(x) \cdot \frac{dx}{dt} + N_M(x) \cdot i_L(t)^2\right\} \cdot N_B(x) + \frac{dt[i_L(t) \cdot N_L(x)]}{dt}$$

Equation 5

This total transfer function for the filter is broken down into non-linear correction functions, which are as follows:

For stiffness:

$$N_k(x) = \frac{[k(x) - k_0] \cdot R_e}{Bl_0}$$

Equation 6

For electromagnetic force:

$$N_M(x) = \frac{R_e}{2Bl_0} \cdot \frac{dL(x)}{dx}$$

Equation 7

For electrodynamic coupling:

$$N_B(x) = \frac{Bl_0}{Bl(x)}$$

Equation 8

For damping:

$$N_D(x) = \frac{Bl(x)^2}{Bl_0} - Bl_0$$

Equation 9

And for inductance:

$$N_L(x) = L(x)$$

Equation 10

The displacement signal required for equation 5, $x(t)$, can be produced from the input signal, $u(t)$, by linear filtering:

$$x(t) = L^{-1}[H_x(s)] * u(t)$$

Equation 11

The transfer function $H_x(s)$ is given by the following equation because both signals obey the linear equation in equation 4:

$$H_x(s) = \frac{Bl_0}{R_e \cdot [J(s) + k_0] + Bl_0^2 \cdot s}$$

Equation 12

However, the input signal for the driver must be produced by a non-linear system:

$$i_L(t) = \left\{L^{-1}[I(s)] * x + \frac{N_S(x)}{R_e}\right\} \cdot N_B(x)$$

Equation 13

Using an additional linear filter with the transfer function:

$$I(s) = \frac{J(s) + k_0}{Bl_0}$$

Equation 14

Together, equations 5, 11 and 13 represent the linear transfer function of the mirror filter in the time domain. Due to the mathematical structure of the transfer function, a filter with such a transfer function could be implemented in a non-linear network composed of discrete elements such as can be seen from figure 8. (Klippel 1992)

Methodology

Design Requirements of the System

As can be seen from previous approaches to reducing low frequency non-linearities that we have explored in the previous section, any system has two main functions. Firstly, the system has to be capable of identifying the parameters of the loudspeaker in its enclosure. And secondly, once the system has identified the parameters of the driver in its enclosure, it needs to use those parameters in order to produce a filter which appropriately reduces the non-linearities of the driver, specifically with the focus being on low frequency non-linearities.

As such, the approach taken by Klippel with his mirror filter appears to provide the most stability with high excursion of the voice coil, as generated by high input signal levels. Being that the non-linearities that this report has focused on in earlier sections are primarily caused by high signal levels, and subsequently higher excursion values, subsequently this system will take on a similar structure to Klippel's mirror filter in order to gain stability at higher signal levels.

With regards to identifying the parameters within the system the system should be a simple as possible from a user perspective such that it requires minimal equipment. By using a microphone and if necessary, taking electrical readings from the terminals of the loudspeaker, by using appropriate test signals to determine the parameters of the driver. This should be done by an iterative method such that it optimises the parameters such as the instrumental variable method as discussed in section 2.3.2 so that the filter is capable of working in the widest range of environments and systems.

The system should also be able to reduce the harmonic distortion of the driver below twice the resonant frequency of the driver by at least 70%. This should be achieved by applying a linearising filter to the input signal such that utilises sufficiently accurately identified parameters of the system. The accuracy to with which these parameters can be identified will therefore determine the efficiency of the filter and therefore the overall reduction in the total harmonic distortion obtain through utilisation of the filter.

The latency introduced by the filter should be no greater than 10ms, and ideally as small as possible in order to make the system as usable as possible across multiple applications.

Development Process

As stated in the previous section, this system had two distinct purposes; identifying system parameters, and utilising those system parameters in order to create an appropriate linearising filter. As such the development process comprised primarily of two parts in line with the two goals.

In order to keep the system as simple as possible, the system was made up of five elements; a measurement microphone, a soundcard, an amplifier, the subwoofer used for testing, and the code created in MatLab which was used to tie everything together. In addition to this a multimeter was used in order to determine the voltage output of the soundcard, and the resistance of the loudspeaker was measured. The voltage at different peak outputs was measured at different digital levels, and the corresponding voltages were recorded. From these points, within MatLab, the values of any digital level's corresponding voltage could be interpolated given that the soundcard's output settings were not adjusted from the measured position.

MatLab was chosen for the environment in which to build this project for several reasons including; MatLab's in-built signal processing capabilities, MatLab's in-built soundcard support, MatLab's graphing capabilities.

The subwoofer that was used was a Visaton WS 25 E – 8-ohm 10-inch speaker housed within a closed box cabinet of 170 litres. This was chosen for two reasons, firstly a loudspeaker of a lower quality with more pronounced non-linearities allows for clearer analysis of filter performance with limited testing equipment. Secondly due to this project taking place in the midst of the ongoing COVID-19 pandemic limited resources were available at the time and subsequently the subwoofer had to be acquired using limited financial resources so the associated costs of a higher performance loudspeaker made them not acquirable for this project.

System Identification

As was mentioned in section 2.3.3 the loudspeaker behaviour could be modelled by using equations 29 and 30:

$$u(t) = Ri(t) + L\frac{di(t)}{dt} + Bl\frac{dx(t)}{dt}$$

Equation 31

$$Bli(t) = m\frac{d^2x(t)}{dt^2} + r\frac{dx(t)}{dt} + \frac{1}{c}x(t)$$

Equation 32

(Knudse and Gru 1989)

Considering that the system under test is only being required to function below 200 Hz, and as was discussed in section 2.3.3 inductance is considered to be negligible for cases where L<<R/2pif and can therefore be approximated as 0. As we can see from figure 11, showing the impedance curve of the driver, and knowing the resistance of the driver from table 1, we can see that L<<R/2pif is true for values less than 200 Hz and are therefore able to approximate L = 0 (Knudse and Gru 1989).

Therefore, with L = 0 being an appropriate approximation for the operating range of the system under test, we can re-evaluate equation 29 as follows:

$$u(t) = Ri(t) + Bl\frac{dx(t)}{dt}$$

Equation 33

Since we know the voltage being output across the terminals of the speaker, the electrical power of the system, and with the resistance of the system being a constant, we can therefore determine the current driving the system at any instant therefore enabling us to identify the first term of Equation 33. As such we're left with two unknowns, the force factor and the velocity of the voice coil, which describe the back emf created through the voice coil moving through the magnetic field of the magnet in the motor structure. Because of the measurements that we are taking with the measurement microphone, monitoring the sound pressure level at any instant, the acceleration of the loudspeaker cone can be calculated as stated by equation 34.

$$p = \frac{\rho S_d}{2\pi r}a$$

Equation 34

From this equation and known parameters about the driver and how the driver behaves in a 3d space the remaining variables other than the mechanical resistance and compliance can be calculated, meaning that these were left to be identified through system identification.

Linearising Filter

Once the parameters of the system had been identified and sufficiently accurately modelled, these parameters were then utilised to create a linearising filter. This was done through use of a filter of similar structure to Klippel's mirror filter, however due to the fact that we had approximated Inductance as L = 0, this meant that the transfer function of the filter took on a slightly different form, as can be seen in equation 35.

$$u_L(t) = \left\{u(t) + N_k(x) \cdot x + N_D(x) \cdot \frac{dx}{dt}\right\} \cdot N_B(x)$$

Equation 35

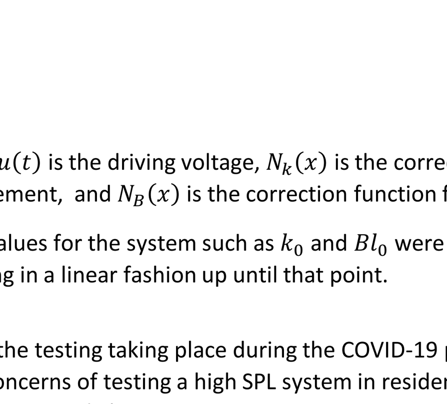
Where $u(t)$ is the driving voltage, $N_k(x)$ is the correction function for stiffness with respect to displacement x, $N_D(x)$ is the correction function for damping with respect to displacement, and $N_B(x)$ is the correction function for electrodynamic coupling. these transfer functions are expressed as follows:

Initial values for the system such as k_0 and Bl_0 were taken to be equal to the values found for the system at the digital amplitude of 0.5, being that the system should still have been behaving in a linear fashion up until that point.

Testing

Due to the testing taking place during the COVID-19 pandemic the system had to be recreated as a software model as opposed to being a hardware and software solution due to noise concerns of testing a high SPL system in residential settings. As such the system was modelled in MatLab which the linearising filter was then applied to with parameters applied as stated above.

Results



Analysis

As can be seen from the results looking at the Total Harmonic distortion of the system, the linearisation filter actually had the inverse effect on the system to what the intended response was at all amplitudes increasing the harmonic distortion within the system.

This could have been due to inaccuracies in the model it self as well as due to a poorly functioning filter as we can see that we would expect the THD to get worse as amplitude increased, however we can see that that isn't the case as shown from our results. As we see the model gets better THD performance around 0.6 before getting worse again until we see an improvement from 0.9 – 1 in the base model. We see signs of improvement in the relative performance of the system at an amplitude of around 0.7 in the filtered signal, however the performance then decreases again. At no point is the performance of the simulated driver improved by the applied filter, but at low amplitude from 0.4 and below we can see that it is not detrimental to the performance. Given that the model was not correlated with any real world data on the loudspeaker behaviour at high amplitude, it is difficult to gauge the accuracy of this data.

Conclusion

In conclusion the system did not have the intended result on the model, as modelled in software however, this could be due to the inaccuracies in the modelling of the loudspeaker itself. For future work having access to testing facilities so that a real loudspeaker rather than an uncalibrated software model would enable for higher accuracy results.

Matthew Monaghan, Q14123550