Sound Systems Design and Engineering

Based on lectures by Dan Perelstein

Notes taken by Daniel Moore

Spring 2020

These notes are not endorsed by the lecturers, and I have modified them (often significantly) after lectures. They are nowhere near accurate representations of what was actually lectured, and in particular, all errors are almost surely mine.

Contents

Part I

Loudspeakers

1 System Design Building Blocks

1.1 Decibel Math

The decibel (dB) involves the use of a logarithm to allow large differences and numbers to be written as conveniently small numbers. A decibel is a ratio between two pressure levels (ie. the measure of deviation of air at a certain point due to sound). These get big so quickly that logarithms quickly become necessary.

There are two definitions for decibels:

$$\text{dB} = 10 \log_{10}\left(\frac{A}{B}\right)$$
$$\text{dB} = 20 \log_{10}\left(\frac{A}{B}\right)$$

The first definition is in reference to power, the second in reference to voltage. In this class, we'll only work with the second one, as that's the one that is generally useful for sound. The first one is useful when talking about power amplifiers. For the future, we'll write log instead of $\log_{10}$ for the sake of simplicity.

A is the pressure value we want to measure, and B is the reference value. The reference value we use (B) is the quietest sound that it's possible to hear.

1.1.1 Logarithms (and their Applications)

A base-10 logarithm like we'll be using answers the question "what exponent does 10 need to be raised to to equal our value?" For example, $10^1 = 10$, so

$$\log(10) = 1$$

This asks the question "what exponent does 10 need to be raised to to equal 10," for which the answer is 1. Another example, $10^0 = 1$, so

$$\log(1) = 0$$

This asks the question "what expoent does 10 need to be raised to to equal 1," for which the answer is zero.

Logarithms convert multiplication into simple addition and division into simple subtraction. Written out, what this means is that we can write $\log(AB)$ instead as

$$\log(AB) = \log(A) + \log(B)$$

Similarly, we can also write $\log(A/B)$ as

$$\log\left(\frac{A}{B}\right) = \log(A) - \log(B)$$

There's another important thing to note about logarithms. If we measure the sound pressure level (SPL) at some point to be x dB, that means that we know

$$x = 20\log\left(\frac{A}{B}\right)$$

How do the number of decibels change for a sound twice as strong?

$$\begin{aligned}20\log\left(2\frac{A}{B}\right) &= 20\log(2) + 20\log\left(\frac{A}{B}\right)\\ &= 20\log(2) + x\\ &= 20(.03) + x\\ &= 6\text{dB} + x\text{dB}\end{aligned}$$

What this means is that any time we double the pressure of the sound, we're adding 6 dB.
Let's see what happens if we do this but we half it instead:

$$\begin{aligned}20\log\left(\frac{1}{2}\frac{A}{B}\right) &= 20\log\left(\frac{A}{2B}\right)\\ &= 20\log\left(\frac{A}{B}\right) - 20\log(2)\\ &= x\text{dB} - 6\text{dB}\end{aligned}$$

So, halving the pressure level means we subtract 6 dB.
One more example: what happens if we have an SPL ten times as large?

$$\begin{aligned}20\log\left(10\frac{A}{B}\right) &= 20\log(10) + 20\log\left(\frac{A}{B}\right)\\ &= 20(1) + x\\ &= 20\text{dB} + x\text{dB}\end{aligned}$$

This means that to get a sound with ten times the pressure, we have to add exactly 20 dB.

1.2 Inverse Square Law

As an introduction, the surface area of a sphere is

$$\text{Surface Area} = 4\pi r^2$$

where r is the radius of a sphere. Note that if we double the radius, there will be 4 times as much surface area.

The inverse square law isn't just a sound thing—it's common between sound, gravitational pull, electric field strength, radiation, etc. If we have a point source that obeys the inverse square law (imagine, for example, a speaker in the air broadcasting sound equally in all directions), the intensity of a sound is inversly proportional to the distance from the source (as distance increases, intensity decreases, and visa-versa). Specifically, they're inversely proportional with a square relationship (hence the name). For example, if we move away from the source a distance of 2m, the intensity drops to 1/4 of what it was. This is because the sound is being projected in a sphere, and as we get farther away, the surface area is increasing, but the total sound pressure isn't increasing, so the sound has to spread out more to cover the area.

As an example. If a speaker 1 ft away from you has volume x. At 2 ft, it would have volume of $\frac{1}{4}x$. At 3 ft, $\frac{1}{9}x$. At 4 ft, $\frac{1}{16}x$. And so on. We get these numbers by squaring the distance and putting it on the bottom.

An important conclusion from this: if you're using surround sound (ie. sound from behind the audience), you'll get more even coverage if you put the speakers farther behind the audience. If the speakers are, for example, 1 ft behind the back of the audience, the volume even 6 feet away (one or 2 rows up) will be already $\frac{1}{36}$ the intensity since it's 6 times as far. On the other hand, if we put the speakers 12 ft away from the last audience member, the volume 6 feet away (one or 2 rows up) is only 1.5 times as far away, so it's only $\frac{4}{9}$ as quiet (if I'm doing my math right). This still isn't the best, but it's a lot better. We can also use vertical height to increase the distance from the audience without having to go as far back.

1.3 Reference Signals

1.3.1 Sine Waves

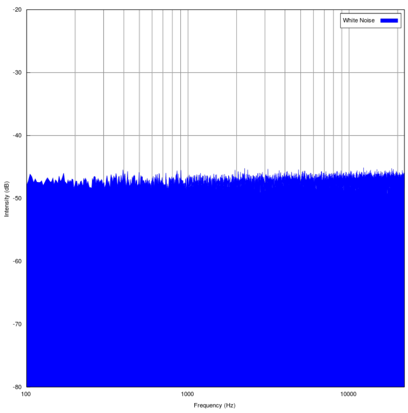
Sine waves are patterns of air pressure where the pressure in some direction can be modelled using a sine curve (very generally, this means that areas of high pressure happen at regular, predictable intervals at regular, predictable maximum intensities, and areas of low pressure similarly happen at regular, predictable intervals and with regular, predictable minimum intensities).

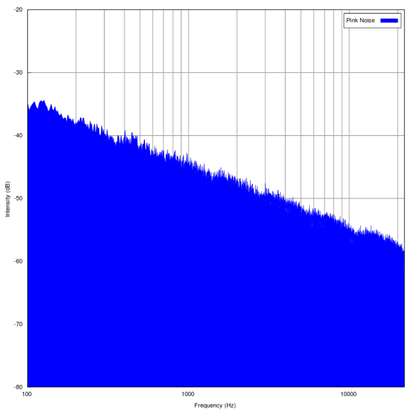
1.3.2 Fourier

Fourier found in the 1820s that any signal, no matter how complicated, can be written as the sume of sine waves at different amplitudes and at different phase relationships. Through Fourier analysis, we can convert from looking at a signal as a function of time, to looking at a signal as a function of frequency. Theoretically, we could totally replicate anything we hear or record with a sum of sine waves, but that would take literally tens of thousands of sine waves every moment, so it's hard.

1.3.3 Noise

There's two types of reference noise we use: white noise and pink noise. White noise has more audible high end than pink noise, and pink noise has more audible low-end. White noise has equal amplitude at every frquency. Pink noise takes white noise and applies a "pinking" filter, which takes away 3 dB per octave. There are more frequencies at a given octave at high frequencies, so in white noise, the high frequencies have much, much more energy, making them overpower the lower octaves. The goal with pink noise is to give us something that sounds like it's evenly distributed betwen notes. Here's the difference between white and pink noise (respectively) on a frequency spectrum:





2 Loudspeakers

2.1 An Introduction to Loudspeakers

Speakers are transducers: electronic components that convert one type of signal to another. Specifically, speakers convert signals from electrical energy to mechanical/acoustic energy.

Ideally, we want to be able to produce every sound that humans can hear (20 Hz to 20 kHz) so we can recreate the sounds we're trying to represent as accurately as possible. This is usually not really possible

2.2 The -6 dB Point

Remember that -6 dB represents halving the sound pressure level (SPL) of a signal. The width that the manufacturers give for a loudspeaker is generally going to e the point at which the volume is 6 dB lower than it is at the same distance on-axis. These levels are measured at 1 kHz, since the high-frequencies attenuate at different rates than low frequencies when you're going off-axis. If we want even coverage, we want to line up the 6 dB point of one loudspeaker with the 6 dB point of the other. If we overlap too much, we'll get areas with more than 0 dB, and if we separate them too much, we'll get areas with less than 6 dB.

3 Vertical Coverage and Tilt

We generally want to range of audience experiences to be within 6 dB, as much as possible. Putting speakers vertically exactly on-axis can make that really difficult if we don't have a very long spacec, because the front audience member very close to the speaker and like 6 audience members back might have a -12 dB difference. There's two things we can do to fix that: first, we can move vertically instead of horizontally. Then, the distance from the nearest audience member and the distance to the farthest audience member is closer to being the same. Similarly, if we have the space, we can try to move it horizontally farther away. That way, the ratio of distance between the closest and farthest audience member is only 1:2 instead of 1:4 or something.

We should try to aim or speakers such that they are pointed at the farthest audience member in the vertical plane. For a raked audience, this usually menas pointing at the top/back row. We want to try to create a situation where gain differences due to axial position and distance compensate for each other.

3.1 Fill Systems

3.1.1 Downfill

3.1.2 Front Fill

If the stage is tall, or downfill speakers would make the imaging very strange (that is, would make a person's voice sound like it's coming from unreasonably high rather than from themselves), we can instead use a front fill speaker or system. Front fill speakers are placed directly in front of the audience, with their axis pointed at the highest relavent audience memeber out of coverage from the main speaker. In a standard theatre, front fill speakers can only cover about the first three rows. We need to take the same frequency response differences into consideration as with downfill speakers.

3.1.3 Delay

What if we're not out of axial coverage, but out of distance coverage? We can extend the coverage of the main loudspeaker by adding another loudspeaker where we lose coverage (at the 6 dB point). We want in this case to aim not necessarily for the objectively highest audience member, but the highest audience member that we're aiming for with that speaker. This additional distance fill speaker is called a delay speaker (or a delay system, if we have multiple). We call it a delay speaker becase we'll have to delay the signal going to it to account for how slowly sound moves through air.
There are some holes in frequency response, so we will have to include both high and low frequencies.

3.1.4 A Rule of Thumb for Number of Speakers

We can figure out roughly how many speakers we need with Bob McCarthy's formula

$$\frac{\text{Distance to the farthest person}}{\text{Distance to the closest person}} - 1$$

This also tracks if we think in terms of decibels: for example, if this formula gives us 1 because of a $2:1$ ratio, that means there's a -6dB difference from the closest to the nearest, so we only need one speaker.

3.1.5 Output Delay

Sound is unique from stuff like light because it's really slow. Light, for example, moves at about $3,000,000$ m/s, whereas sound only travels at 340.29 m/s, or roughly $1,000$ ft/s, or 1 ft/ms. In the time that it would take a beam of light to go around the Earth, a sound would only go about 150 feet.

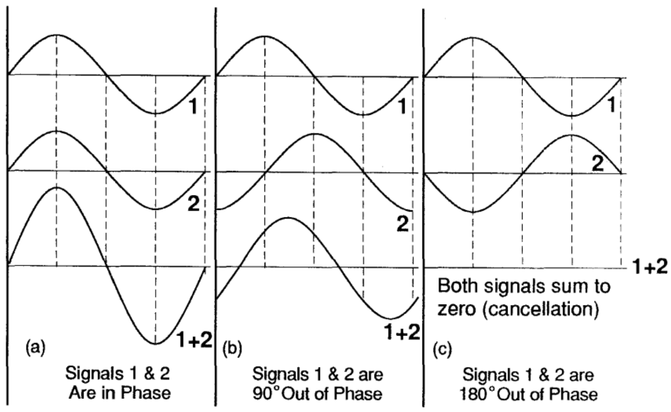
Especially when we're talking about delay speakers or systems, we need to pay attention to this: we need to add delay equal to the time it takes for the sound from the main speaker to reach the delay speaker.

4 Comb-Filtering: Coverage in the Horizontal Plane

4.1 Phasing/Comb-Filtering

When you hear the same sound from 2 different loudspeakers, or you pick up the sound from two different microphones, it sounds bad in unpredictable ways. As much as possible, we want our sounds to be heard from only one loudspeaker, and we want our sounds to be picked up by only one microphone. We'll be mostly talking about loudspeakers right now.

When two sine waves interact, the resultant combined sine wave is equal to the sume of the two sine waves. We can see how this happens below: if the waves come in at the same time, they'll boost each other by double in every place. If they're coming in 180° out of phase, which is half a cycle behind, they'll cancel each other out exacty. Phase is related to the time that the two waves come in relative to each other (that is, 90° out of phase, means that the second wave comes in 1/4 of a cycle later than the first one).



When two sounds are added, we get interference. There are two types of interference:

(i) Constructuve interference: the two waves are combining so that they're adding to each other—the resultant wave amplitude in that location is greater than either of the waves individually

(ii) Destructive interference: the two waves are combining so that they're subtracting from each other—the resultant wave amplitude in that location is smaller than one or both of the waves individually

Maximum constructive/destructive interference happens when the waves double each other or totally cancel each other out (respectively).

We can also apply this to pink noise: when we delay one pink noise signal compared to another, it forces the resultant pink noise to have nodes (places where certain frequencies are decreased). The farther apart we set the pink noises, the more nodes we have and the lower they go, since lower frequencies require a farther distance to be totally out of phase. This *also* also applies to sounds in general: voices, music, etc. At around 30-40ms later, we hear it as a delay rather than a drequency response issue.

For loudspeakers, the "problem path" is where there's little difference in volume, but noticeable time offsets, because this is what causes more intense comb filtering. We prefer level offsets much more over time offsets, and level offsets can help compensate for time offsets a little bit. This means we want to minimize the region of overlap, because what's where we run into issues of comb filtering (and echo, potentially).

With microphones, we talk about the 3:1 rule: however far your first microphone is from a source, your second source should be at least 3x that distance.

4.2 Focus in Plan: Pan

True stereo isn't really possible for more than a couple poeople at a time down the center line of the speakers. It also gets less exciting the farther back you go—very far back, the angular width of stereo differences gets pretty small, and it's not really that exciting any more. For the people outside of this range, a change in pan just winds up being a level change.

Bob McCarthy gives a rule, which we can use as a starting point but doesn't always work, for how to place loudspeakers. He says to divide the audience in two, then point the loudspeakers at the center of each of these halves.

Important: in the horizontal plane especially, parts of your system *will* overlap in coverage. The important thing to consider is the relationship between level and arrival time: they should either be as similar as possible, or as different as possible.

4.3 Fill Systems in Plan: Extending Horizontal Coverage

Like in the vertical plane, we figure out how to point the speakers for maximum coverage (and minimum overlap), and add fill speakers where there is no coverage to fill in the difference. Like with vertical fill systems, horizontal fill systems are named pretty self-explanatorily, like center fill, side fill, etc.

Part II

Mix Engineering

5 Analog Mix Engineering

5.1 Bus Architecture

A bus refers to console outputs. Buses are organizational tools that collect audio inputs on route to a common destination. For example, if we want to send a mix to the stage manager that's separate from the main mix, we could say that those mixes are on two different buses. Buses are used in analog consoles, digital consoles, recording studios, live sound, DAWs, etc. Although this is in the analog console section, it's not specific to analog consoles, this is just the first time it was convenient to bring it up.

There are a few basic types of busses, which we'll go over in more detail in the following sections:

- Aux (or variable send)
- Group (or fixed send)
- Stereo/mono (or master)
- Matrix

5.1.1 Aux Bus

Aux busses, also known as mix busses, are variable sends—this means that you can control how much sound gets sent to this bus individually, as a function of the fader and the aux send. Aux busses can only receive input channels. Aux busses are usually used for effects sends (eg. reverb) or monitors.

5.1.2 Group Bus

Group busses are fixed sends—this means that you can't independently control how much goes to the group bus, only whether or not sound is going to it. The volume of the channel going to the group is just set by the fader. Group busses can also only receive input channels. Group busses are typically used to group channels together that are headed towards other output busses. A good example is grouping up all the drum mics before sending them to the main output.

5.1.3 Master Bus

Also fixed sends, and also controlled by the main fader. Master busses can receive any combination of input channels and output busses, unlike the last two. When the term "stereo bus" is used to refer to these, it's ued in a generic sense of the term (ie. two speakers), not the literal definition of stereo.

5.1.4 Matrix

Matrices are variable send busses, but matrices can receive only some combintion of output busses. The levels in the sends are a function of matrix sends and the fader level. Matrices are basically aux busses, but for other busses rather than input channels. These are typically used to combine output busses to fill systems.

5.1.5 Pre-Fade vs. Post-Fade

This is only relavent for aux sends and matrix sends. These sends can be pre-fade or post-fade. Pre-fade busses mean that the audio is diverted to the aux bus before it's sent to the fader, so both the fader and the send get the same input (ie. the send volume doesn't rely on the fader—the main mix will have no effect on the bus). Post-fade busses have audio diverted after the fader, so the input for the send is the same as the output of the fader (ie. the send volume also relies on the fader—changing the main mix will change the volume going to the bus as well).

5.2 Analog Consoles

6 Additional Considerations for the Engineer

6.1 Analog Patching

6.2 RF Basics

7 Stages of Amplification

7.1 Power Amplifiers

7.2 Gain Structure

8 Paperwork

9 Microphones

9.1 How Microphones Work

9.2 Selectio and Deployment of Microphones

9.3 Feedback

9.4 A:D and D:A Conversion

10 Digital Mix Engineering

10.1 DCA Assignments

10.2 Digital Consoles

11 Delaying Larger Systems

11.1 Input Delays

11.2 Delay Matrixing

11.3 Staging Zones