



Technical Support

Simple Tom  v1.8

Room and monitoring

Referencing and listening levels

Metering

Common mixing issues

Excessive low end

Too much reverb

Mono-compatibility issues

Frequency buildups or resonant frequencies

Combined negative characteristics

Dips in the spectrum

Harshness

Unbalanced levels for instruments in the same class

Over-compression due to chasing loudness

Defining mix translation

General mix considerations

Tonal balance and texture

Clarity and definition

Dimension and depth

Emotion

Mixing decisions

Source sounds and arrangement

Gain staging

Levels

The Equal-Loudness Contours

Identifying resonances

The high mids

The low mids

The low end and mixing bass

The low end hierarchy

Mixing vocals

Processing the vocals

Preparing the lead vocals

Preparing the backing vocals

Vocal de-essing philosophy

Time-based effects

Dimension and Depth

Stereo optimization

The 3D virtual stage

Saturation

Compression

Compression and psychoacoustics (masking)

Parallel compression

Upward compression

Multiband dynamics

Limiting

The master bus chain

Why use a chain on the mix bus

Using the same, mostly “set it and forget it”, master bus chain

Building the master bus chain

Good practice

File management

Exporting the project - pre-roll and fade-out

Sample rate

Bit depth

Dither

Audio file formats

The end

Room and monitoring

Our brain is very capable to adapt to almost any space (within limits), and apply internal filtering to make everything sound natural. When we review our own video snippets from live parties, we always notice lots of reverberation in the video, while we know for sure the sound was great. That’s the doing of the brain behind the scene.

And that’s also the reason why we need to take ear breaks often. When things go out of balance, the brain is “helping” us hear in a clearer manner our flawed mix, and in the process we go even deeper on the rabbit hole of an unbalanced mix.

If you can’t work in an acoustically-treated space, you can still achieve a good listening experience in a **very narrow sweet spot** (exactly in your usual position at the desk), by using

a correction method (Sonarworks, Arc etc.). This, coupled with a pair of corrected headphones that you hear properly (we hear uniquely different, chose what works for you, not by brand or hype), no doubt will help you reach “radio-ready” mixes.

Just remember, mixing as a skill takes time to develop.

We will always be biased by our own taste, so make sure your monitoring allows to hear everything in a reference mix. And make sure you constantly improve your taste. The best working context is when our taste and our critical listening overlap almost perfectly - so we don't need to make that switch between being “creative” and being “critical”. We can be both, at the same time.

Maybe visit a bunch of showrooms with auditioning rooms, and studio control rooms, play your references there and see if they reveal additional details. If you do hear important details there, and not hear them at your place, investigate. See how you can mitigate the issue.

Referencing and listening levels

I recommend using good references to reset your ears to a neutral overall tonal balance, and not to copy individual elements. I always recommend Serban Ghenea. His mixes are absolutely next level. Regardless of your genre, if you can get a similar tonal balance, you're in the big league. Of course, use a similar volume when referencing. If you find yourself routinely pushing your monitoring levels hotter and hotter as the time passes, try this: during ear breaks, play a reference at a lower level, so you'd get used again with lower listening levels. Then you can resume working on your mix at a lower level. High, sustained levels are tiring regardless of how we work - monitors or headphones.

When switching back from a reference to your mix, notice what you feel: is your mix darker? Then you need to push the elements responsible for the high end, more. Is the other way around? Do the opposite.

Especially if we use too deep narrow bell cuts in our tracks, it will come out fast when we compare to references. The references sound clean and open, while a mix with too deep narrow cuts will sound in a very unique “twisted” way which I can't describe. For this reason, I recommend to do some EQ practice where you do such cuts at various frequencies, only to get familiar with the sensation. It happened to me so many times in my early mixes, and I never managed back then to find the cause, because the mixes were otherwise balanced, but the tracks themselves were missing chunks of information and that lack was affecting the mix audibly.

Listen (to your mix and the references aswell) at medium/comfortable levels for overall balance, loud levels especially for the low end power (and potentially harshness and

sibilance), and quiet levels for clarity and density check - make sure all elements can be reasonably heard even at low listening levels.

If your space allows, position yourself out of the direct sound coming from the monitors (side of the desk, or alternatively under the desk). The low end response is different enough to provide extra cues regarding the mix.

Metering

Except for the *correlation* metering and low end measurement with the *spectrogram* (discussed in the “*Common mixing issues*” section), metering is mostly informative. We can easily hear a 3 dB boost or cut at any frequency, but on the FFT analyzer, even if it is a slow one, it’s close to impossible to visually see the difference. So if we hear that something is wrong, we can turn to metering to see if it confirms, or to pin down the exact point where the issue resides. If we can’t see anything, we act on solely based on our ears.

Try to not get extremely hooked on metering. Always remember that metering is also related to dynamics. When you don’t really understand why two references have different metering readings (for the spectrum tonal balance) and yet they seem like having a similar sound, that’s why: they have different spectrums and dynamics control, but in both cases the combination of spectral balance and dynamics results in a similar overall sound. This is taste in action: the engineers made the moves that sound right (regardless of what the metering showed).

A paid solution for referencing and extensive metering is Metric A/B placed on the mix bus, the very last plugin in the chain. In addition to metering and isolating different bands, it provides a way to quickly switch between your mix and a list of references. You can map the A/B function to a keyboard key or MIDI controller, and make the switch easy and natural.

Otherwise, you can always place few references (inline, in a single track) in the work in progress project, and just solo that track when you want to preview the reference. You may have to make some adjustments to prevent the reference passing through the master bus chain, but it’s fairly easy (use a summing bus dedicated to the reference, and one bus dedicated to the project, and place your master bus chain there).



ADPTR Metric A/B

Common mixing issues

Excessive low end

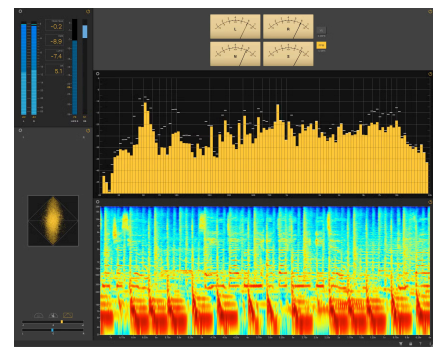
Small speakers (and a taste for nice, fat, sound) will eventually bias us into adding **excessive low end** to our mixes. Make sure you have monitors that can properly reproduce down to 30 Hz. You'll need a fairly big SPL at this frequency. I can't spend more than few minutes with this kind of low end, so I do my work on the kick and bass, then turn off the subs (or activate a 6 dB/oct gentle filter on the master channel).

And then, I won't do any low end adjustments, and work on other parts of the spectrum. Every time I work on the low end, I just turn on the full spectrum monitoring.

A frequent side-effect of excessive low end is distortion when listening through small consumer speakers (or smart phone speakers). The strategy here is to really keep the low end under control (10 dB of compressor gain reduction for a consistent bass is not uncommon), monitor on the smartphone, and observe a RMS ballpark at which the smartphone is not distorting the low end, and keeping that value in mind for the next mix.

For metering, using a reliable spectrogram solution (like [IK Multimedia Metering](#)) is mandatory. There is no better way to investigate the low end as a spectrogram: orange = potentially too weak, red = ok, spots of dark-red = ok for kicks, but when it comes to basses it may be too much. Continuous traces of dark red = way too much.

When I see continuous traces of dark red, that means I push the low end frequencies too much.



IK Multimedia TR5 Metering

Why? Because I don't hear them properly. Why? Because they're not properly represented in the upper frequencies (too quiet or not enough harmonics or both) and I'm essentially pushing that low fundamental into distorting it with the limiter. This is my sign that I have to re-visit the low end, re-shape the kick, add something else to the bass (a cabinet emulation, saturation, Waves MaxxBass, something to make it richer) - then I won't need to push them that much.

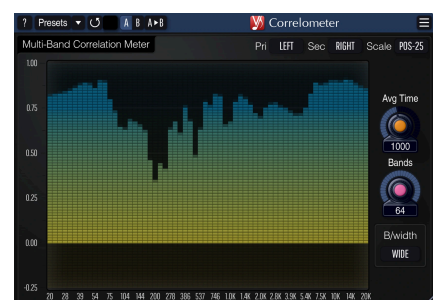
Too much reverb

The room and speaker placement will also influence our choice of reverb level. **Too much reverb** can often happen. The rule with the reverb is to push it until we hear it, the dial back few dB. Then mute it and un-mute it. We should almost not hear it when we unmute it, but we should definitely hear it missing when we mute it (like the sound gets smaller). Make sure you're properly filtering what gets out of the reverb. Low-mids tend to muddy up the mix quite fast.

By the way, the reverbs won't add more frequency content, so whatever you send, the same frequency content gets out of the reverb. So EQ-ing before or after reverb leads to identical results.

Mono-compatibility issues

And this leads us to potential **phase problems** when collapsing the mix to mono. Many synth built-in effects, in order to make a spectacular stereo sound, neglect the phantom center (mono-compatibility). Just a handful of these, or a suboptimal choice for the main reverbs in the project, can completely throw off the stereo image of a mix, leading to lots of phase problems. Normally we can hear phasing, but if you don't, just use a (free)



Voxengo Correlometer

correlation meter, and keep the mix in check for phase issues.

Frequency buildups or resonant frequencies

They pop-up on metering rarely, most of the time they don't. You need to hear them, and act on them on a track-by-track basis, not by slapping an EQ on the master channel. Many times a resonance can be the main character of a certain important sound, so if this is the case, just make room for it in the mix, by massaging down that frequency range from the other tracks.

Combined negative characteristics

Mixing harsh elements (hats, shakers, overly loud and uncontrolled vocals) with dull elements - probably in an early attempt to make everything sound better together, can often result in a mix exhibiting these effects at the same time: harsh AND dull. Make sure every new element sounds good in the mix. Don't leave it for later, you'll get used to it and then bias the entire mix.

Dips in the spectrum

Aggressive (narrow) EQ cuts in an attempt to fix issues can lead to big **dips in the spectrum**, which can lead to a special kind of "frustrated" and weak, unbalanced sound. You need to practice EQ-ing sounds, to be able to recognize this effect, because the spectrum of the entire mix can look alright, but those deep cuts over various elements will add up and impart their sound. We always need to EQ out the least we can get away with. Soothe is rarely the answer. A strategic combination of EQ (up to a couple of cuts) and compression works better every time.

Harshness

Especially the high-hats (everyone is using the classic drum machines) have multiple **harsh frequencies between 1.5-6 kHz**. If you have them sit right, the entire mix will sit right in the high end, so make sure you spend time sculpting them (and even better, save your own kit). De-ess, transient design, a bit of saturation, whatever it takes to manage those spikes.

Unbalanced levels for instruments in the same class

The kick and bass should roughly have the same loudness. The click of the kick, and the top of the clap. That "tick-thlap" alternation between the kick and (kick +) clap when you listen to the high mid range. Vocals and piano. And so on. It's fine to have some louder than other, but there is a limit on volume difference from which the sound becomes too "all over the place".

Over-compression due to chasing loudness

Over-compression, clipping, over-saturation on individual tracks or submix groups (the infamous clip-to-zero thingy that gets more and more popular) will quickly suck the life out of a mix. Adding density (and by extension, loudness) to a mix requires great amount of skill and experience. Squashing transients is definitely not the way.

Defining mix translation

Translation means ensuring that all - or the most important - elements of the mix are well represented regardless of the listening device. Mono or stereo playback systems, small or big.

I'm seeing in the last years something I call "*expensive cheap sound*": mixes made on expensive sound systems that sound pretty lifeless on average consumer devices (or over the ambient sound systems in stores). By relying exclusively on very expensive monitoring, the engineers have more precision in placing the elements very targeted, and they forget to properly represent the kick and bass into the upper frequencies (low midrange) or the high hats into lower frequencies (midrange). So what they get is a mix that sounds great on €20k+ speakers, but not so great on €20 speakers, where there are not enough audio cues for the low end and high end, so a lot of movement and groove get lost.

At the opposite way of the spectrum are, you guessed, the engineers that still rely on speakers that reflect the consumer-grade equipment reality, like [Serban Ghenea](#), or and another mixing engineer with excellent taste here in Europe, [Monte](#) - he mixes all Purple Disco Machine releases. Serban Ghenea uses just ProAC Studio 100 (with a subwoofer), and Monte a combination of Auratone cubes, Yamaha NS-10 and a pair of more accurate midfields. They get absolutely stellar and characterful mixes for consumer devices, which sound equally great on big sound systems.

Funny enough, [Hannes Bieger](#) (next in line in my preferences after Monte, mixes for Stephan Bodzin and others), after a complete studio upgrade with a fantastic set of Genelec in-wall farfields, one month later added a pair of ProAc Studio 100 because he probably realized what I just explained.

That's not to say cheap monitors are the key to great mixes. A pair of ProAc costs over \$3K. But we need "down to Earth" speakers, that represent the reality of many listening devices, while also being able to reveal mistakes (and, of course, to cover the entire spectrum from 30 Hz to 20 kHz).



ProAc Studio 100



Avantone MixCubes

General mix considerations

How to optimize the mix. The mix is always about balance. Within a track, between tracks, between sections. The main technical aspects we have to always keep in check are tonal balance and texture, clarity and definition, dimension and depth - all working together to convey (or enhance) the story of the song and ultimately, the emotion.

Tonal balance and texture

Probably the biggest enemy of translation: masking. It can manifest within a track (a certain frequency area way louder than the rest). It can manifest between the tracks - louder frequency areas in a track will mask how we perceive the tonal balance of other tracks.

The expensive speakers can deal with way more excess of any frequency than regular speakers. Depending on how well you know your monitors, you may or you may not realize what they reveal. Spend lots of time listening music with them, so you become intimately familiar with all their quirks.

Make sure all elements keep roughly the same (perceived) levels. That they are not competing too much (unless you do a deliberate choice to place certain elements at a higher level).

Then check the representation I mentioned earlier: is the deep bass having at least a couple of top harmonics so it tickles the midrange? If the kick is low and thumpy, does it have a nice complementing midrange “tock” to represent it, or at least a slight click? Are the super hi-fi hi-hats which you placed at 11-14k, represented at 4-8k aswell? This kind of questions. Of

course, it's impossible to fit the entire mix in the midrange, but little hints for all the elements are perfectly possible, and they make for a great translation.

The wider elements will move at the back of the mix when listening in mono (so you can make the secondary elements wider, and the main elements narrower, so in mono you still have all the important elements of the mix). The tonal balance of the Mid and Side should be roughly the same, so the mono down mix will be perceived similarly as the stereo mix.

Clarity and definition

We should be able to identify every individual component of the mix. The upper mid-range should be very well controlled (by controlling every track of the mix, and additionally slightly massaging the mix bus with multiband compression). The transient content of the mix should be consistent across all tracks, so they are separated enough but not “too separated”, at which point the mix becomes difficult to enjoy.

Recordings and vocals must be cleaned of any unwanted noise, clicks, pops, sibilance. Especially for drum kits recorded with multiple microphones, you need to spend time and make sure that everything is time-aligned. Same for background vocals.

I had a “phase” when I was obsessing about clarity and definition, my good friend and music producer warned me that my mixes are too clean, boringly clean, and I should dial it down a few notches. I protested a bit, but eventually I realized he was right. I suppose it was a natural step. After going raw (because I couldn't hear better), I went to the other extreme - very clean, then back to a more balanced approach. If this happens to you, I think it can be a sign that you made significant progress - you have all the knowledge to make things squeaky-clean, you just need to work on the taste to bring the balance to a more “digestible” (natural) level.

Dimension and depth

I wrote on the topic of creating the virtual stage quite in depth (pun intended). If a reverb is very important for a vocal, consider making it almost centered, or use a pair of reverbs (one wide and one narrow). The idea being, that even when listened in mono, the mix should still carry some of the most important elements that set the stage.

Then, at this point I suppose it is obvious, all the main elements should be very well represented in the center of the stage (kick, bass, leads, vocals).

Stereo elements, the wider they are, the more are reduced in level when listening in mono (with as much as 6dB for some, and 3 dB for hard-panned elements). You can use this to your advantage, to create two layers - one for the stereo mix, and one for the mono mix

Emotion

Make sure the character of the song is still present regardless of the listening medium. It takes an equal amount of skill to convey a sonic identity be it of a clean or of a dirty mix, or raw or smooth. Make sure the character is there. That I can feel the polish or the rawness even through a smartphone speaker. And by extension, I can feel the emotion of the song. If you manage to place the emotion in the spotlight (the vocal, the instruments, and all the elements that are helping build it), you have the winning formula.

Mixing decisions

Source sounds and arrangement

Mixing starts with the quality of the source sounds. The best mix comes from quality recordings, good samples, good synth presets (without a lot of mud or excessive distortion). Distorted sounds will lose transients, therefore definition. It takes a lot of skill to get a punchy and distorted sound (with strategic parallel distortion). The choice of sounds will determine how far you can take the mix.

Next, is the arrangement. A rock band sounds very full with just a drum kit, bass guitar, lead guitar, rhythm guitar and vocalist. In dance music, we tend to overdo with layers for synths, percussion loops and transitions or sound effects. Stephan Bodzin's projects are usually under 20 tracks, sometimes even a bit over 10 (excluding sends).

Every new element will take resources (spectrum, space inside the virtual stage), make it count! If it's redundant (takes the same space and role as an existing sound), pick the best one of those two.

Then consider the flow of the arrangement. Proper contrast between sections (verse and chorus/drop) leads to a better perceived mix. We need segments with less energy, less space requirements, in order to perceive the chorus/drop as being massive. If everything is wide, nothing is wide. If everything is massive all the time, nothing is massive. Think how a rollercoaster plays with emotion and anticipation.

Gain staging

I won't make it more complicated than it needs to be: gain staging consists in making sure each plugin in a chain receives the optimal signal. The optimal signal is at the very least under 0 dB FS, and for some analog emulations is around -20 dB FS (otherwise they saturate or even distort).

From my experience, the optimal choice is to place a gain plugin first on each track and bring everything down to peaking at -20 dB FS. From there, no matter what processing follows (even 10-15 dB EQ boosts), its very hard to bring the output of the track over 0 dB FS. So this utility trim gives me peace of mind that most of the time I'll have no gain staging issues.

Levels

Using a master bus chain (as detailed in a different chapter), will already impose limits and help diagnosing for issues. Also will force us to get whatever we need from our mix (dynamic range, density, perceived loudness) as we go along, because there will not be any later processing that can radically change things.

Set the main anchor(s). Usually at the beginning of the project we don't know exactly how the top of the kick will be perceived in the context of the mix. Have an EQ ready, because we usually need to make boosts or cuts to give the kick that clarity it needs, but without being too intense.

A trick to quickly set the volume of any track in the mix is to pull the volume down to a point where I mostly don't hear the sound in the mix, then I boost 3-4 dB. This is usually a good ballpark to start with, and will keep adjusting it as I progress with the upcoming processing moves.

We instinctually set the levels for sounds according to their most prominent midrange spectrum area. We may then notice that other areas are needed (for example the low midrange sits right, but the sound in the mix is dull, unclear and undefined). Depending on the context, a simple EQ boost may work for bringing life into the sound. Or a bit of saturation. Or compression, to let the initial transient be louder, and then bring the body of the sound down to not crowd the mix.

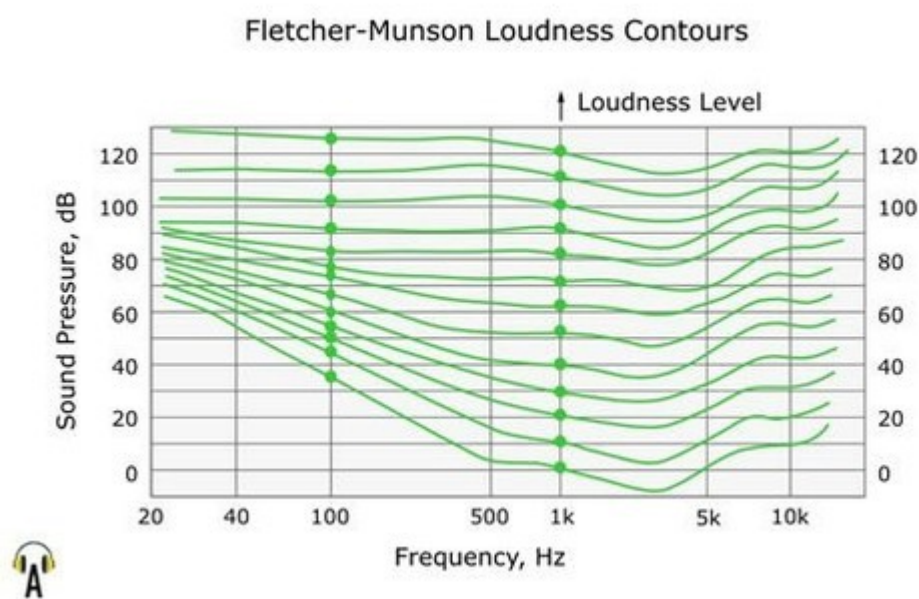
Our ear is sensitive around 3 kHz area (again, the Equal-Loudness Contours) if there is too much content (or resonant content) in this range, the sound will feel harsh. We usually instinctively dial back, to avoid harshness. The problem is, in the exact same area resides the clarity of a sound. So we really need to find a way to have a consistent level for that sound in this range, without sounding harsh. Most of the time the transients are the ones that exhibit the harshness, so de-essing, (multiband or single-band) transient shaping, multiband compression, are all valid choices in an attempt to mitigate harshness.

I think you realize at this point that writing a full book on mixing, and trying to cover every situation and every strategy, is mission impossible. For every given sound in every unique context, there are a multitude of solutions we can try, and there will not be one best. Each decision we make, will influence the rest. All we can do is to gather experience by making

hundreds of mixes, and then use our experience to make good decisions during our current mix.

The Equal-Loudness Contours

These contours were made with the help of human participants. They were given sine test tones at various levels and frequencies, and their perception of loudness was then processed into these curves. Some say that because the test tones were sines (and not with transients), they are potentially flawed. We can, however, still extract very useful insights from them though.



The Equal-Loudness contours

How to interpret this series of graphs: Say I play a 1 kHz sine test sound that reads a bit over 60 dB SPL (7th green dot from the bottom). This is my reference loudness perception.

If I were to play a 100 Hz sine, I follow the green line starting from that 7th dot, and I notice that the dot corresponding to 100 Hz is way higher than 60 dB, it's almost 80 dB. That means, the ear is so insensitive to low frequencies, I have to play them almost 20 dB louder to perceive them as loud as 1 kHz.

A 100 Hz sine played at 80 dB will be perceived as loud as a 1 kHz sine played at 60 dB.

If I go the other way, I can see that around 3-4 kHz the ear is more sensitive and a little under 60 dB SPL at these frequencies is perceived as loud as 60+ dB at 1 kHz.

At the other extreme (the highs) the ears become less sensitive again, and we need to boost them if we want them to be perceived equally loud as our 60 dB 1 kHz sine.

Now what are the practical implications of this? If we listen an old 80s song (Michael Jackson - Bad - the original master) at very low volume, we hear almost no low end (and sometimes almost no high end) and we mostly hear the mid range. Now if you keep the same volume and play a recent mix (as always, Serban Ghenea never disappoints), you will surely notice that you can hear almost the entire spectrum from the low end all the way up to the highs.

What changed? The density of the mix. The ear doesn't perceive audio as "loud" only when it's played loud, but also when it's compressed.

One important aspect of this phenomenon is this: a mix that was made with no master bus processing and sounds good unmastered, will receive some EQ treatment at mastering exactly because a different density requires a different tonal balance for the mix to be perceived as tonally balanced as it was initially. This is one reason for using a master bus chain from the very beginning of the mix: the chain will add density to the mix, and the tonal decisions will be more informed.

Also one reason mastering engineers often mention using De-Essers or other resonance suppressors in their mastering chains: they receive very dynamic mixes (that are correct in their respective state), but mastering density is getting things out of their pockets. If you use your master bus chain, you can properly De-Ess whatever needs De-Essing, and not the entire mix. Even if you remove your chain before mastering, you will send a mix that's already tailored for density, so the mastering engineer can do a way better job.

Identifying resonances

Resonances and dominant frequencies have pretty much similar masking effect. First of all, it's not mandatory to suppress them unless it's either obvious it's masking the track or it's masking other elements in the mix. Often times they give character to a sound. Take the cymbals in the TR-909 machine drum. If we notch out the 3-4 kHz resonances, we strip the identity of that sound. You can try doing this suppression, or boosting the higher end, to mask the resonance with a more prominent high end. Pick whatever sounds better for your mix. The

idea is that we can always have the option to boost around a resonance, not only to suppress it. It's a delicate operation to work on these, take your time.

You can use the analyzer of the EQ to spot for POTENTIAL resonances. The easiest is to set the analyzer slope to 0 dB / octave. That makes the frequency content to be displayed lower in level as the frequency increases. So if we spot frequency buildups bigger in amplitude than the amplitude of the lower frequencies, there is a reason to investigate. It may be nothing, though. Vocals will often exhibit this behavior with no actual sonic issues.



FabFilter Pro-Q 3 spectrum slope setting



FabFilter Pro-Q 3 potential resonance display

It's really on a case by case basis. If you do the "sweeping technique" you'll start finding resonances everywhere, you'll start removing them and the resulting sound will be completely lifeless. This is what Soothe usually does: trying to conform the incoming audio to a pre-set internal slope, completely disregarding the actual sound. Doing small narrow cuts left and right rarely sounds good. You really need a compelling reason to take such measures.

The high mids

The high mids around 3 kHz (let's say 1.5-5.5 kHz) are the "presence" area. The ear is most sensitive at this frequency. Most of the sounds in the mix, including basses, will have some representative harmonics in this zone. This is relevant for when the listener is using small speakers so they can still get the perception for all the sounds in the mix.

This area is a bit of a double-trouble: when we have too much here, we hear really clear all the elements with harmonic content here, but an excess in this area will also mask the perception of the rest of the mix (in both directions: high and low end).

So it's crucial to carefully control all elements that can crowd this area (vocals, instruments, cymbals, kick top, snare, clap - really a lot of them). As soon as this area is under control (when the harshness is removed), the "air" usually reveals itself and there is little to do (in terms of a bit of excitation and EQ) to make it nice and apparent.

So for example if you are struggling to add “air” to your mix, often times the move is not to boost highs, but to control the high-mids. If we just boost the high end of the mix, we're pushing the harshness aswell. How many times you had a mix sounding harsh and dull at the same time?

At the same time, if we cut too much in this area, we lose the clarity for those sounds.

And it gets a even more complicated. The Equal-Loudness Curves show us that our perception of spectrum balance changes with the listening level.

In our mix, is the final density (loudness) that dictates how the ear perceives the mix. And this happens regardless if the mix is actually listened to at loud levels. The simple fact that we reduce the dynamic range of the mix (we made it dense and louder) already dictates how the ear will perceive it at any listening level.

Here's the relevant link between these curves and the high-mids: the denser our final mix, the more we have to control the 3 kHz range (usually compress downward), to avoid this range (now boosted by the dense mix) to overpower everything else.



Dense (loud mix) = likely the high mids will need a cut. Dynamic (less loud) mix = high mids can potentially need a boost for clarity.

So the amount of this midrange area we need to adjust, will depend of the mix density (loudness). A more dynamic mix will need less midrange control. A very dense and loud mix will need more careful and precise midrange control.

It may be the case, as the mix progresses and things get louder, to go back and revisit this mid range on many elements, until the mix falls back in place.

This is one of the many reasons I use the top-down technique (a master bus chain): I can listen to my mix at the final density and I can make clarity adjustments already knowing how much the mix will be pushed.

The low mids

The low mids (150-500 Hz) are the real glue of the mix. Not bus compression. We (our ear) have a special relationship with the low mids thickness. Not enough low mids, may initially sound cleaner, but the mix will suffer from lack of weight. Too much low mids will quickly mask everything in the mix. The optimal level is an open but strong feel for the elements in the mix.

When in doubt, place a -30 narrow EQ cut at 150 Hz. If the mix opens up, then you probably need to address the low mids more carefully. If the mix gets really clear but light, then you are good (or very close).

200 Hz and below is primarily reserved for the low end, maybe the bottom of a snare, and the bottom of a vocal. There is some power and richness in the vocal in the 100-200 Hz range so it can't be cut entirely, otherwise the vocal will be weak sounding. Often times a 3 dB low shelf cut at 115 Hz can leave just enough low end for the character, but will cut enough to avoid clashes.

300 Hz is where the lowest synths/other low melodic elements retain their energy. Therefore a bit of kick and bass scoop in this area may be useful (since these two are better represented lower in the spectrum - and also higher, as basses often get a boost between 700-1500 Hz, for that extra presence). Also, for sub bass definition, this area must not overtake the subs in terms of energy. The vocals can be scooped as well in this range.

400-700 Hz can get muddy quite fast - but adds power to many elements. Scooping this area will push the elements at the back of the mix, helping in adding depth. On the other hand, this low mids area affects the way we perceive the upper 1-6 kHz area (the perception of smoothness and harshness). A too scooped range in the vocal for example, can make it appear harsh even though it was not harsh in the first place. Leaving this area untouched on many elements can make the area too prominent. Scooping everything can add separation and clarity but also increases the perceived harshness.

Each of these ranges must be assessed one by one for each track and decide the priorities. Every dB counts.

Finally, the 300-500 Hz range (due to higher harmonic density) tends to contribute more on triggering the limiter. Worth spending extra time and make sure the limiter is not triggered excessively (or exclusively) by this area.

As usual, my recommendation is to go heavy handed in cutting these areas for learning purposes. Once you hear what can go wrong, will be easier to avoid in upcoming projects.

The low end and mixing bass

Below there are the tables showing the low end octaves (1 and 2) notes and frequencies. There is no reason for anything under 32 Hz to be prominent in the mix (the energy radiation is best seen on the spectrogram). If a kick radiates energy way under it's pressure point (typically 48-55 Hz) investigate why. If a bass does the same, investigate why. Other sounds that have no reason to have low end should have absolutely no reasons to radiate energy around or below 30 Hz.

C1	32 Hz
C# / Db	34 Hz
D	36 Hz
D# / Eb	38 Hz
E	41 Hz
F	43 Hz
F# / Gb	46 Hz
G	48 Hz
G# / Ab	51 Hz
A	55 Hz
A# / Bb	58 Hz
B1	61 Hz

C2	65 Hz
C# / Db	69 Hz
D	73 Hz
D# / Eb	77 Hz
E	82 Hz
F	87 Hz
F# / Gb	92 Hz
G	97 Hz
G# / Ab	103 Hz
A	110 Hz
A# / Bb	116 Hz
B2	123 Hz

For the non-bass elements, I'd cut with a 12 dB/oct filter at around 70-75 Hz.

For the kick and bass, I'd investigate exactly what produces that low frequency under 30 Hz. Sometimes it may be an oscillator tuned very low. We can't remove it, because it's part of the texture. But maybe we can remove the fundamental. If that's not possible, we try filtering - 6, 12, 24 dB/oct. whatever works best.

The idea is to individually apply the optimal filter for all the elements that have low end issues, to the point that there's no need to do it on the mix bus. On individual tracks we have the luxury to fine tune the cut frequency for the sound. On the mix bus, we're kind of forced to compromise.

If you have issues in balancing the lowest octaves notes, just play them individually from the keyboard, and check if one of them is getting too prominent, or one of them is too quiet. You

can then alter the velocity of the notes in the piano roll, to bring all the notes in the bass line at a similar level. We can always compress the rogue low end harmonics with a targeted dynamic EQ, but it's better to fix at the source as much as possible.

There are not too many harmonics in a bass. For D note the fundamental is 36 Hz, then the 2nd, 3rd, and 4th harmonics are 72 Hz, 108 Hz, 144 Hz. We can really manage the behavior of the low end of a bass quite granular. And we can follow the behavior of each harmonic, when checking for issues.

The midrange is a bit different. Too much midrange will affect how “deep” we perceive the bass. Monitoring is important. Or spectrogram, headphones and ear breaks every 5 minutes. The low end is numbing our ears quite fast.

There are all kinds of basses. In drum'n'bass or dubstep, where there's a lot of gliding, dynamic EQ may be the best bet for catching resonances. The rule is simple: if you hear the resonance every note, then use a static EQ. If you hear the resonance here and there, go with dynamic EQ.

Once the sub is set in relation with the kick (at most equal as the low end of the kick, most of the time, 3 dB under), if you want more or less bass, work on the upper area. Do not change the low end! Investigate what you can boost to get more perceived bass (maybe 1-3 kHz, maybe a bit of saturation etc.). The idea is that the low end can't go any higher, so you need to work on other parts of the spectrum that still can be worked on.

Sometimes the solution for more bass is something simple. After you've set the current bass with proper low end and dynamics in the low end, just send it to a parallel bus with over compression and an amp simulator. Bring in some extra parallel midrange and suddenly the bass comes to life.

The low end hierarchy

The low end, especially the lowest octave, is really not permissive in terms of how many elements can be present at the same time. I came out with a simple ducking solution, to establish which elements are allowed to add their low end into the mix, and it goes as follows. The kick has maximum priority.

I am sending sidechain signals to all the groups in the mix (Bass, Transitions, Percussion, Instruments, Vocals, Returns) and I am ducking their low end with a dynamic EQ low shelf (-20 dB).

I am then sending another round of sidechain signals from the bass to all the remaining groups (Transitions, Percussion, Instruments, Vocals, Returns).

Every time the kick hits, the entire project's low end gets ducked. Every time the bass hits, the entire project's low end gets ducked. When these two are not active, the rest of the project can send low end information and that will not be ducked. This method allows to fill the spectrum at all times (sometimes with the kick and bass, other times with the rest of the mix), and doesn't sound like the lows are always scooped out from the other elements.

Mixing vocals

In reality we start the vocals processing with RX cleaning. I start this chapter with the typical vocal chain because it will often bring out things that we neglected initially. On the other hand, we also have a pretty standard formula to work with vocals. So it's a bit of back and forth.

Processing the vocals

So I'd do a first pass of cleaning the stuff that needs 100% removal, but not spend too much time on details that I won't completely remove (but volume ride), like breaths or types of mouth noises needed for the performance.

The classic formula for vocal processing is:

Pitch Correction (Melodyne, AutoTune etc.): this goes before any other kind of (audio plugin) processing. These voice correction plugins work best on untouched material, especially because they occasionally introduce their own artifacts. The entire processing that comes after, should also take these artifacts into account. Use them to taste.

Compression: the 1176 keeps the intelligibility of the vocal brightness, aiming for between 3 dB and 6 dB of compression. 10 dB for the "in your face" sound.

EQ: generally roll off the low end and boost the highs; high pass filter about 75 Hz, high shelf adding about 3 dB at about 5 kHz to taste.

De-Esser: squash back the most prominent sibilance starting around 5.5 kHz, 6-10 dB. If more de-essing is needed, going manually via volume is the solution, as opposed to setting the de-esser any harsher.

As more unwanted things will poke through, they should be written down for RX editing. At this stage we also have a good ballpark for the volume of the breaths (how much they're over what we want, and how much we should edit, so when they pass through the vocal chain they sound as we want them).

During the progress of the song, we do volume rides (as opposite to more compression, which will sound unnatural and “processed”) over phrases, words, or even syllables, to make the vocal sit in the mix (not above the mix), and make the performance sound natural and powerful.

Preparing the lead vocals

The vocals can make or break a good song, a good production, and a good mix. iZotope RX is your friend when it comes to addressing prominent mouth noises (spit clicks and pops, smacking lips, mouth opening sounds, breaths, pops, esses). Some need straight up removal (clicks, spit noises, and the sound of the mouth opening), but others make for the character of the voice, so they need volume riding (or EQ riding for pops).

End-of phrase exhale breaths when removed will hint that the singer’s performance was effortless. Perfectly fine to do when the song is high energy, maybe not the best idea to remove them when it’s a more intimate kind of song. It’s always about finding a balance where it sounds that the singer is not struggling but not faking it either.

A short inhale (volume ride and fade) again, sounds natural. If the singer is intended to sound like they were whispering in your ear, that might lend itself to letting you hear some more mouth noise.

In some instances, such as a totally quiet drop in the music where the vocal is going to come in alone, that mouth opening or licking-the-lips sound might be meaningful and authentic. Each noise and breath needs to be evaluated in context.

It might be too unnatural at that point to not have a soft breath or other close mic noise. Similarly, before a big sustained note, that breath might add power to the performance; the singer is really going to go for it here.

There is no automatic solution for cleaning vocals like this. Each noise removal, fade in, fade out, or volume ride needs to be listened to and evaluated as to how it helps or doesn’t help the performance.

Preparing the backing vocals

In the backing vocals, the problem is just as critical. The good news is that we can go more aggressive with the removal of unwanted elements (cut out all breaths and mouth noises - especially because being backing vocals they’re often out of time). We have a lead vocal in the spotlight, though.

When backing voices end their notes at different times, I’ll often fade them all to end together. I generally want clean and precise backing vocals on every note, as if they are sung

by a choir, and not by one singer doing multiple takes, and also to avoid the build-up of headphone bleed across the multiple takes.

Vocal de-essing philosophy

It needs to sound natural but not distracting. Too much de-essing will sound like the singer is lisping, not enough will be painfully obvious at loud volume. Backing vocals can be heavily de-essed and again, even cut out if the timing between vocals is off.

Reverbs and delays may point out places where noises or breaths become distracting again. All mouth noises, and breaths may need to be taken out of delay sends, because they can muddy up the sound.

Soothe. For me this is an audio recovery tool, not a mixing tool. I manage to stay away of it, most of the time.

Time-based effects

For vocals, a handful or carefully-chosen effects can go a long way. The real skill is to automate “throw sends” (works great with delays), just certain expressions, words or even letters. That will accentuate that part. Furthermore, using different throws on the same (repeated) chorus, will make it appear different, and keep the listener engaged.

A typical start-up selection would look like this: reverb 3 second, like a cathedral, reverb 1.8 second, like a wood room, reverb 1 second, like plate or algorithmic, chorus, quarter delay, eighth delay, dotted eighth delay, and half delay.

When dealing with a more intimate context (vocals need to appear really close to the listener), a very small room, or even just early reflections would probably be a suitable choice.

Dimension and Depth

Stereo optimization

Samples from sample packs, synth effects (especially Nexus), even some reverbs, can sound superb in stereo but completely collapse 6 dB lower when listened in mono. That means they don’t provide a solid phantom center. It’s a good idea to narrow all tracks (or even mono them), and then use your own mono-compatible send effects to restore the width and depth. It will make a world of difference. The stereo field issues are some of the hardest to fix, if not taken care of soon enough.

The 3D virtual stage

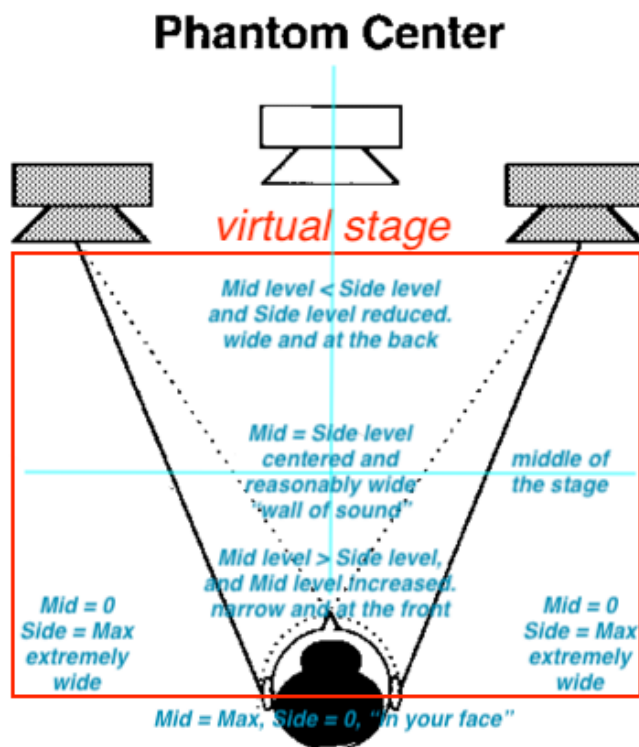
The stereo speakers are creating a 3D virtual stage for the listener

First, some listening for context: face a pair of speakers, music playing, eyes closed.

Although there's just two speakers, you hear sounds appearing from left side, or right side, or any position in-between. Some elements (like the vocal) appear to come right from the center. But the speakers are positioned left and right, and there's no speaker in front of you. Let's see why this happens.

There are two main aspects to understand:

1. Those two speakers are creating a 3D virtual stage that has width (Left to Right), depth (Back to Front) and height (Bottom to Top). The mix you're listening to, will have various elements spread inside this virtual stage: for example, backing vocals or pads would be at back of the stage, spread from left to right. The lead vocals on the other hand, will either dominate the center of the stage, either center and exhibit some Left-to-Right spread.
2. When we mix, our main aim is to arrange all the elements of the song into this virtual stage, by manipulating what signals we send to the left and right speakers, and what processing we apply to these signals.



The concept of Phantom Center: a perceived sound source that appears to come from a central location between the left and right speakers, even though no actual speaker is positioned there.

This illusion is created by playing the same audio signal at equal levels and at the same time from both the left and right speakers. In case of a reference mix, the Phantom Center and the Real Center (the middle line between the speakers) of the stereo field, are perfectly aligned.

Things change when we pan individual stereo sounds to intermediary positions, but not to extreme Left or Right. In the case of let's say a 50% Left panned stereo sound, the Phantom Center of that sound would be a line coming slightly from the left (not quite as extreme as the black dotted line, but from the same direction).

The Phantom Center is what we call Mid, in the Mid-Side technique.

Mid represents what's common between L and R channels. Side represents what's different between L and R channels.

If you grab a Mid-Side balancer and start adjusting the Side level of the reference mix, if you reduce the Side, you'll end up with the entire mix narrowed down in the center of the stage. If you do the opposite, the mix will sound very wide, like you moved all the players at the extreme left and right of the stage, while the perception of the center will decrease to the point of having no perceived center anymore.

What Mid-Side processing can do to a sound depends on the available levels (since we can't go over 0 dB FS). I'm starting front to back.

If the Side level is at the maximum the mix allows, and the Mid level is smaller, then the perception of the sound will appear to be from outside of the speakers - coming from extreme left and right, lacking a defined center.

Equal maximum levels of Mid and Side will place the sound in the "in your face" line, but being wide, it may not be perceived as focused as the next one.

Maximum Mid level and smaller Side level (or zero Side level) - now the sound is still "in your face", but this time it's progressively focused in the middle of the stage as the Side level decreases.

For all these 3 above, we need really good reasons to place them at the very edge of the stage. This space should be reserved for a main element (lead or vocal).

Smaller than maximum, equal side Mid and Side levels: that would be the average levels for most sounds of the mix, in order to ensure a solid "wall of sound" at the middle of the

stage. Small variations in Mid and Side levels will bring the sound forward or at the back (Mid level) or make it slightly wider or narrower (Side level in relation to Mid).

Lower levels will place the elements at the back of the mix, and again their width will be dictated by the Side level in relation to Mid.

Now when it comes to the mix as a whole, in order to have a phantom center, a focus zone, the Side level can't be louder than the Mid level.

As soon as the entire mix has a Side level bigger than the Mid level, we are starting to experience that “split” in the center of the mix, and everything starts to lose focus.

When the phantom center is non-existent, the mix doesn't really have direction. It's like all elements are staying at the sides of the stage and no one has the courage to step in the spotlight. Everything sounds wide, but undefined, and we can't really identify these elements' position inside the stage.

This is why imaging on the master channel is a bit of hit and miss, if the mix didn't mindfully controlled the width and panning of each element.

Mono compatibility

And now we circle back at the title: an easy way to check for the phantom center is, you guessed, checking the mix in mono.

We don't actually care about the mix sounding good in mono - the good old “mono compatibility”. But we care about the stereo mix having a well-defined spotlight: the phantom center. It just happens that this mono-compatibility check is an easy way to discover issues in the virtual stage. If the mono information turns to be lower in level for important elements, then we basically leave important real estate unused in the virtual stage. And we really want to fill up that stage as much as possible. Or if not fill it, we still want to place some elements in focus.

Purely mixing in mono doesn't make too much sense. I mean, if you want to get good at fitting elements inside a virtual stage, you should definitely mix in stereo, listen to each sound, and decide its placement by the means of L-R panning, Mid-Side panning (for stereo sounds), EQ (less highs will place the sound both lower in height and push it towards the back of the mix), dynamics (less transients aka less definition will again push the sound towards the back of the mix), room reverb and pre-delay (the bigger the pre-delay the bigger

the perceived space, but at the same time more pre-delay will place the sound closer to us, while less pre-delay or zero will place the sound furthest).

Depth in modern mixes

Modern music, in a typical chorus, due to less dynamic range, doesn't exhibit too much depth. It works more like a wall of sound, where all the elements are positioned at the same depth, and arranged (for separation) left to right and bottom to top. Of course, the lead, kick and vocal will be centered, sometimes very close to the "in your face" point.

Then, the verses are a bit more relaxed and there can be more stage depth here.

I am mentioning this because if you want to make your music fit a certain established genre, it's a good idea to follow their rules.

The "in your face" character of lots of genres can only be obtained by aiming towards a dense mix. Having too much depth, although may sound nice, has the potential to place that production outside of the target genre.

I reiterate, I don't say less or more depth is better, I just say that is a good idea to know your genres, and make informed decisions when it comes to depth, width and height of the stage.

Saturation

Saturation can enrich a sound with the cost of slightly losing definition (the peaks will get rounded) or can destroy a sound. In my opinion, we can get away with less saturation than people tend to believe is needed. Vocals, pianos, other instruments that have a relatively clean spectral content during sustained notes, will not handle saturation too nicely.

We usually get saturation during sound design, during compression, even during limiting on the mix bus. A typical use for saturation is on basses, to add harmonics. But even then, maybe a bass amp and cabinet simulator work better. In short, saturation is definitely not the magic mix savior, "character, width & depth" provider people pretend it is. It's just another tool in the box, one very unstable for that matter (because it's highly level-dependent).

In addition to adding harmonic content (density), saturation acts like EQ (by adding a boost at the frequency range where the harmonics reside), and also like compression (by shaping the amplitude of the source signal usually downward). Given this complex interaction with the sound, its utility is something you truly need to try for yourself. Saturators will increase the level (even when they do nothing - when the input signal is not enough to trigger saturation) so you need a solution that can do a proper gain leveling for before and after. iZotope Ozone Exciter does it really good. During your technique trainings, try it on everything at very small percentage and full-blown distortion. Hear it and then remember where you can use it.

Compression

I want you to **compress everything**. Especially if you don't hear compression easily. If you don't use compression, if you don't overdo it for a while, you will never actually train your ears to hear it. So use it, overdo it, then you'll start hearing it and dial the right amount. Worth sacrificing a bunch of songs, for learning how compression sounds, because it is such an important tool.

If you have no idea where to start, a digital compressor with 16ms attack and 160ms release, ratio 2:1 or 3:1 and 5 dB of gain reduction would be a good start. After the meter clearly shows that it reaches 5 dB of gain reduction, use the compressor's built-in make-up gain to increase the output with 5 dB. Additionally, measure the peaks with and without compression and adjust the make-up until you read the same value. Now turn the compressor on and off. With no changes in level between the before and after, what you're left to hear is compression. If you can't hear it, maybe change sources: claps, snares, kicks, hats, plucky bass lines, plucky synths.

The reality is, in today's landscape, most mixes need compression to achieve the expected density. Although people recommend not to use compression if you can't hear it, and I routinely discard any processing I can't hear, my suggestion is to use about 3 dB of compression on all tracks in the project, even if you don't hear it. Make-up the level for each compressor. Then either (depending on your DAW) map a key to on/off all compressors at one key press, either export two separate versions of the mix: one with compression, and one without, and compare if you hear a difference in the context of the entire mix. It definitely should be obvious. If you can't hear it, just practice more on hearing compression. Go more extreme with the settings: 6-12 dB of gain reduction. Try various monitoring levels, until you find a combination that allows you to hear how the sounds get shaped by the compression. You really need to hear it, to really progress to the "next level".

Live recorded drums can have a huge dynamic range (prominent transients). If the starting-up compression settings can't deal with these transients, you can bring the attack towards 1 ms or even use a limiter to make the work of the compressor easier.

The material coming from sample packs may be already compressed, apply compression anyway, listen how it sounds, and only keep it if it sounds good.

The rest of the instruments and vocals, they all can deal with the start-up settings. However, in case of vocals, they are a combination of plucky sounds and sustained sounds. It's often common to use a combination of a limiter to tame the peaks (like the 1176 emulation) and then a slower compressor to stabilize the sustained parts (like the LA-2A emulations). Compress until you get character, then volume automation if some words are not popping out clearly. Over compressing vocals sounds really "unprofessional".

Compression can't always help (often doesn't) in keeping a sound in the spotlight (clarity-wise). So be prepared to employ level automation throughout the arrangement, to keep in focus the elements that really need the spotlight at that point in time.

Compression and psychoacoustics (masking)

If you want to have a deeper understanding of the auditory masking, there are two videos that explore the phenomenon:

<https://www.youtube.com/watch?v=9gSsK9AzviM>

<https://www.youtube.com/watch?v=sKNyyVILl3o>

Otherwise, you can skip them and take my word for these two statements:

- The amount of time a sound A can be louder than another sound B, without A masking B, is 20 ms.
- If a sound A masks the sound B (A louder than B), when the sound A stops, it takes about 200 ms for the ear to clearly hear the sound B.

When we mix, we want to hear all tracks, that's a perfectly normal target to have. In doing so, if we have several tracks playing at the same volume and at the same time, masking between them can occur.

There is another phenomenon happening, independent of the masking. If the transient of a sound is strong enough, and then the body gets ducked by compression (or ADSR), the ear will "reconstruct" the body. We can get away with less level for that instrument. This means better mix quality (better separation), and more headroom to get a loud mix.

Obviously, what we can do, is to compress the tracks, leave their transients pop, and then duck the body. Because not all tracks will have all their transients happening at the same time, the ear will pick up the transients as they pop, and do the identification trick.

This is a fine balancing act, as it can easily lead into over-compression. But it's a natural step if you want to have better mixes.

Think a piano riff playing with vocals. Uncompressed, the resonant piano notes will most likely mask the vocals. With compression, just the transient of the piano notes will be loud, allowing our brain to ID the sound, and then the vocals will sit nicely over the ducked piano tails.

In a balanced mix, this is how we get that extra separation and definition: leaving as many transients as possible a bit above the "body" of the mix.

If masking happens, our auditory system will recover (and perceive mix definition again) 200 ms later after the masker ends. In a mix, chances are that masking takes place repeatedly, and adding those additional 200 ms will result in a permanent masking effect.

It's important to prevent this phenomenon, by typical means: gain, pan, eq, phase, dynamics. Usually, by applying a compression strategy that lets transients pass through, many other masking issues will be avoided. When it eventually happens, we deal with it on a case by case basis.

What follows is a bit of coincidence. The attack of a compressor represents a speed (at which the ratio gradually increases and reaches the set ratio), and not a delay (for when compression kicks in).

Most compressors have their default settings at around 10-20 ms attack. Usually this default is the value that allows those up to 20ms long transients, and avoids masking by clamping down on the body of the incoming audio.

Now, this is just one use case of compression. It doesn't mean all the time compression should be set like this. But for transient-enhancing, it's a good start (the 15-20ms attack value).

Also, doesn't mean that all tracks need compression. Just use your judgement and experience, and when you clearly deal with separation issues and you can't turn the elements louder anymore, start thinking about de-masking strategy with dynamics manipulation.

Parallel compression

There are two things you can accentuate with parallel compression: transients, or low level information (bringing up more details). For transients you need to set heavy compression with slow attack and medium release, and for the latter you need a faster attack and medium-fast release.

Why going parallel and not using the dry-wet knob available on many dynamic processors? Because a 90% dry-wet knob will reduce the dry with 10%, and I don't want that. I want to keep the exact set level of my dry signal, and just add extra something else.

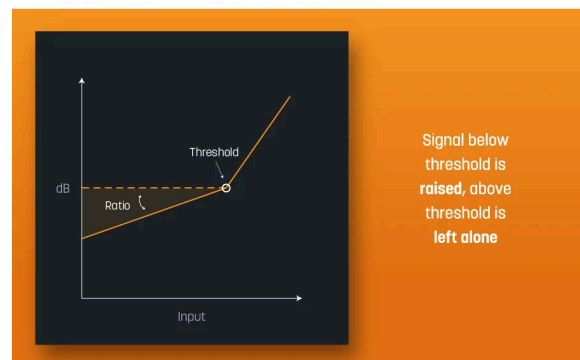
Speaking of parallel, there's a cool trick we can do: say we have all instruments going each to a parallel compression bus to bring up details (so smashing everything). When the vocal is playing, send a higher level from the vocal to that compressor. It will react by outputting a lower level (higher input level, more compression, lower output). But because it also receives

level from instruments, so it will output *less* level for the instruments. Essentially the vocal is directly influencing the density of the instruments, making them slightly less denser while the vocals play. This “direct ducking” sounds way more natural than a typical spectral sidechain (people use separators like TrackSpacer), where the filters will introduce phasing artifacts.

Upward compression

All of the above uses the conventional downward compression. There is another, and arguably cleaner way of revealing the details of a sound: (multiband) **upward compression**. One thing though: the attack and release settings will sound exactly the opposite of conventional compression. So if we want more transients with upward compression, we set the release (not the attack) to a value like 20ms. If we want no transients from upward compression, we set the release to 0ms. Ratio up to 0,8:1, threshold as high as 0, adjust to taste.

UPWARD COMPRESSION



This is a very elegant way of adding rich details and texture to our mixes, especially in the mid-range and highs, especially with a multiband plugin like Ozone Dynamics, which allows us to blend different proportions for each band.

A quirk of upward compression: it lifts the noise floor. As you can see from the transfer function above, the gain starts small at the threshold and increases as the level of the input signal decreases. We can end up with uselessly large amounts of gain at -90 dB. Yes, with this one iZotope didn't do enough in-depth thinking. And they did the same mistake with newly-implemented feature of upward compression in version 11 of Ozone Maximizer. We have other options though. Good ones.



Ranged upward compression (red) versus non-ranged upward compression in Ozone Dynamics (white)

In the screenshot we have a really useable and good sounding upward compression transfer function: range 2dB, threshold -1.5 dB, max ratio - as in the final screenshot below. It can be easily obtained with Blue Cat's Dynamics 4 (red transfer function).

In white, we have the typical upward compression (no range, just ratio) implemented in Ozone Dynamics, Ozone Maximizer 11, Sonnox Oxford Limiter (the Enhance function).

There are two issues with the white one. First, the audible boost happens lower than our interest area.

The 2 dB boost is applied to levels under 8 dB, and for a modern, snappy sound, we really need the ability to dial in upward compression immediately under -1 dB.

The second issue is with the lack of range: eventually the low-level information (think -40 to -70 dB) will get a higher than 2 dB of boost. We can mitigate this with a noise gate, but this is too much work.

For general use (non master bus uses), a very gentle lift with this type of no range upward compression works good enough. But when it comes to precision, when dealing with a dense signal on the master bus, where each dB matters, the ranged dynamics becomes a mandatory tool.

For wide band processing, Blue Cat's Dynamics 4 works great. For multiband, simply hosting multiple Dynamics instances inside Blue Cat's MB-7 multiband mixer also works great.

This particular setting will leave intact everything above the -6 dB threshold, and will gradually add up to 5 dB to everything under the threshold.



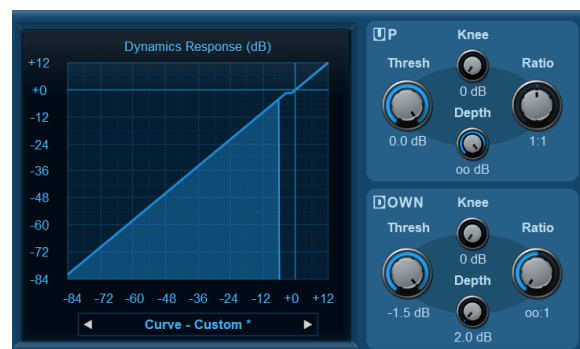
Blue Cat's Dynamics 4

If you use this technique on the master bus chain, especially with this much gain in a multiband setup, make sure you're mixing into it from the beginning. It will lift your reverbs

and other low-level details, so if you place it after finishing the mix, it will make them few dBs louder.

Now instead of clipping, you can add before or after the limiter a small amount of upward compression (see the next screenshot). It will add more density and details, while completely preserving all transients.

A little goes a long way, even 1-2 dB can have a positive impact (clearly audible) on the final mix, so make sure you're not overdoing it.



Blue Cat's Dynamics 4 mastering setting: adding 2 dB of extra density without touching the peaks

There are several powerful techniques out there - to achieve loud, yet crystal clear, punchy, with a rich low end, mixes. These techniques only work when used from the beginning of the mixing process, because they completely change how we are balancing the elements of the mix.

Be aware, they can do the exact opposite and destroy a mix when used just at the end, slapped like a band-aid. This is one of them.

Multiband dynamics

I'm using it mostly as part of the master bus chain, or when a certain frequency range on a track needs a fix that can't be done with a static EQ or, thinking at vocals, de-esser. In the past I used the low band of a multiband compressor to duck just the low end of the bass keyed from the kick. Nowadays I use a dynamic EQ low shelf for the task (namely Sonnox Oxford Dynamic EQ), I can hear (and I visually confirm it on the oscilloscope) that it does the attack and release faster. Often times I just smash the kick and bass into the mixbus chain (namely the multiband) and let it deal with both, and it often can do it perfectly fine.

In a mastering context (so when processing a stereo file), multiband dynamics is rarely a solution. It can alter the tonal balance of the mix, and this decision is not for the mastering engineer to take - it's a production decision.

When mixing, multiband dynamics can be a powerful tool (as previously described), when used from the beginning of the mixing process.

Excessive clipping or limiting will not result in a clear, defined, punchy and loud mix. It's very easy to have a loud mix, but at the expense of the low end punch, clarity and definition. By using multiband dynamics

In the case of a song with a typical punchy kick and a heavy bass (can be pop, trap, dance, electronic), a well-made mix will have the bass about 5-6 dB under the kick. This ensures a proper contrast (distinct kick thump and bass oomph) while allowing for a really loud mix. The low band of a multiband compressor can keep them under control and maintain this contrast. A clipper won't. A clipper will just shave the kick (because the bass will rarely to never hit the 0 dB FS ceiling), and as a result will reduce the contrast between the kick and bass. The thump will be reduced and the low end will be transformed into a continuous "hurrrr". I often use Ozone Maximizer with a very fast character (to emulate clipping in a way), but I prepare the levels for the kick, bass and the rest of the mix in such a way, that the clipper will really deal with just random, rogue peaks. It makes no sense to limit the kick, but it will limit when the rest of the mix plays on top of the kick and random peaks pop up.

As for treating buildups on the master channel, I watch the limiter action, and when it's clear that is fighting to fix the mix, I go back to the mix and fix the sources of the buildup (including by automating EQ during the chorus, where many tracks play), than letting the limiter do the job, because a limited sound is losing definition very fast.

Limiting

I think the best way to view the limiter is as a very fast level rider in relation with the 0 dB FS ceiling.

If you zoom-in around the point of the highest the peaks of the waveform, the limiter would assess every individual peak in that area, and decide how to deal with it. A limiter must have 0 attack to avoid letting clips slip through, but it can get quite complex when it comes to the shape and timing of the release, to avoid distortion while preserving loudness and transients.

The less it works, the more it changes just tiny fractions of the highest peaks, and the limiting process becomes more transparent.

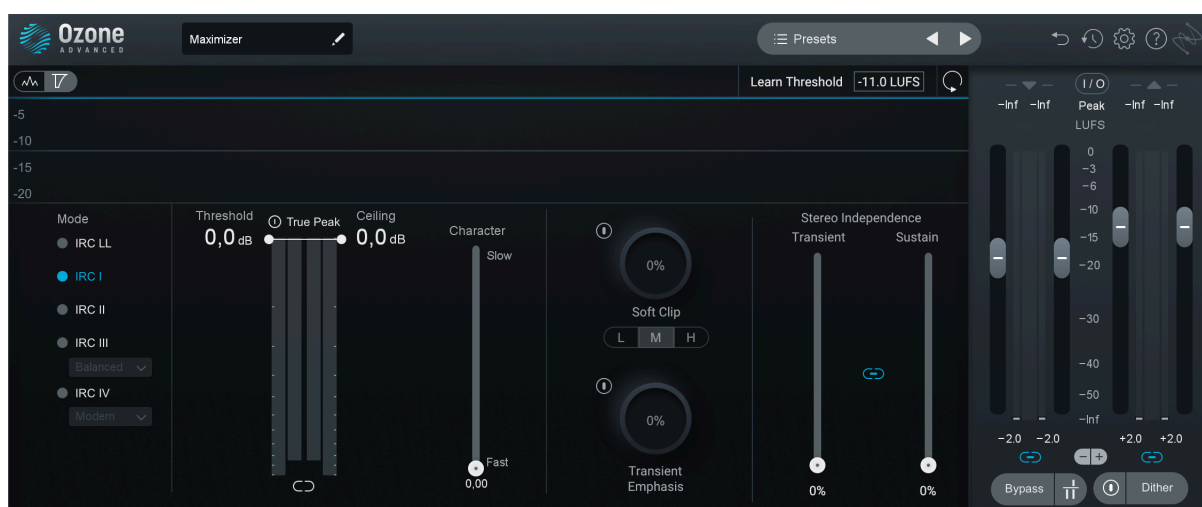
The shortest distortion always hides behind the actual transient, and we can't really hear it. This is why I recommend using two limiters in series. In the case of two limiters, the first will take just the top half of the peak to be limited, while the second one will work only on the bottom half.

Yes, this means that the first limiter will rarely work (1-2 catches over a 4-bar, for example) but it matters. You won't need to aggressively clip all the tracks, or busses, and still have a loud, and especially dynamic-sounding master.

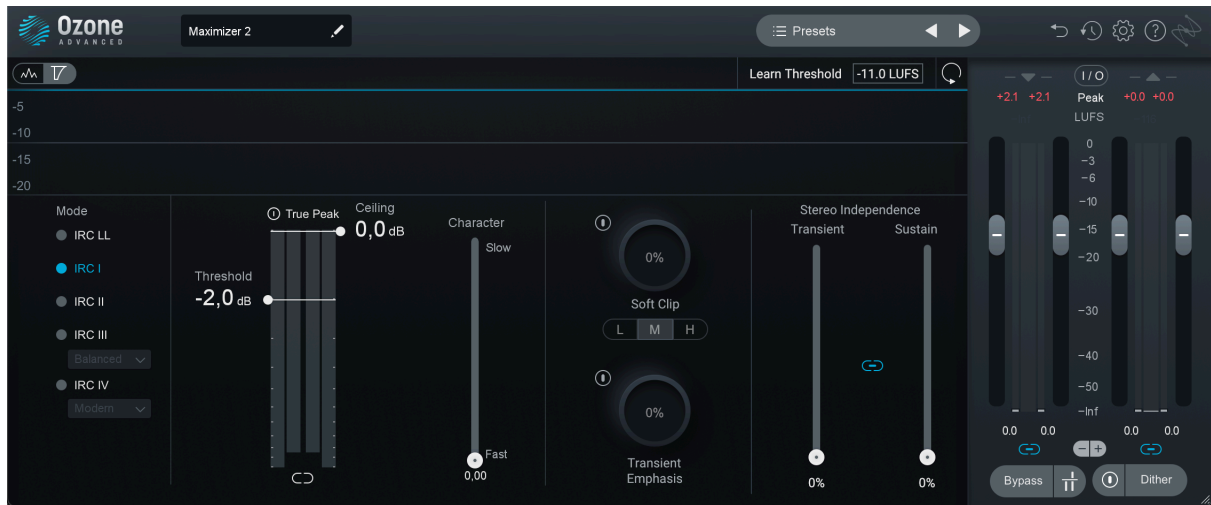
Dual-limiter configuration

The “pre-limiter” will have the input gain set at -2 dB, and the output gain at +2 dB. What this means: any signal entering this limiter with an amplitude up to +2 dB FS will not be touched, because the input is reduced with -2 dB. All +2 dB signals become 0 dB. Any limiter can be used, even if they don't have in/out controls, by simply placing two gain plugins before and after it.

However, let's say a 3 dB peak is hitting the limiter. Subtract -2 dB (input) and the actual signal entering the limiter becomes +1 dB. This one will trigger 1 dB of limiting, and will output a signal of +2 dB. Of course, the final limiter must deal with the excess, but it will deal with just 2 dB instead of 3 dB. If the peak is very short, this entire limiting process will be virtually inaudible, due to spreading the job across two processors.



Pre-limiter: the input gain is set to -2 and the output is set to +2



Final limiter: used as usual.

Now, why a pre-limiter and not a clipper? I feel like clippers distort too much. The limiter tries to ride the level, and doing it in stages makes a difference in sound quality. Also many limiters can actually be driven into clipping behavior - in Ozone case, with the Character slider.

Speaking of limiters, do ISP (Inter-Sample Peaks) matter? I can't hear them. Many main big mixing/mastering engineers do not use ISP protection, and their mixes have thousands of ISP's across a song. Other big mixing/mastering engineers do use ISP protection. Do what you feel. It's not that important, definitely not to the point of making or breaking your master.

The master bus chain

Why use a chain on the mix bus

I think that one of my top 1 “a-ha moments” was when I realized how important is to hear the mix in progress at a “mastered” level. The saying is “make a mix so good, it doesn't need mastering”. Yes, I agree. But we can't make such mixes without having a proper chain on the master channel. Every mixing engineer who's mixes I love, does that (I later found out). A master bus chain, mostly “set it and forget it”, from the very start of the mix. If you produce, add the chain as soon as you finished recording, and you don't need zero latency anymore.

Why “master bus chain” and not “mix bus chain”? To remind me that I am using a mastering engineer mindset: gentle touches, clean saturation options, “hi-fi” moves, generally speaking.

I can't stress enough the importance of the master bus chain. Our perception of music (tonal balance, transients, separation) is highly level-dependent. As the Equal-Loudness Contours describe it. We don't make the same mixing decisions (EQ, leveling) when using no compression (not a single compressor on the entire project) compared to when we use compression (every track has some form of dynamic treatment).

A denser mix (using a master bus chain) represents the tracks in a different way than a very dynamic mix.

The point I am trying to make is that the density of the mix is vital! If we want to avoid guesswork as much as possible, **only** when we work within the density that we expect for the mix to have in real life (post-mastering) we can make proper decisions for every dB or half dB of automation, EQ, compression etc.

If you ever tried to master mixes that you've made with an empty mix bus, you probably noticed that new details popped out at mastering. Some were pleasant and many were not. And you didn't address them at the mixing level for a simple reason: you couldn't hear them. But after mastering, the mix often becomes denser with 6 dB or more, and now you hear those lower level details better.

We can only address the tiniest details when we can hear them. So we need a master bus chain to provide us with that density that comes after mastering. Now, the sound of this chain will become part of the mix. That's ok. Most of the time you won't need mastering. Your mix is your master. If you want further mastering, just disable the final limiters, export the mix in 32-bit floating point no dither, and send that file for mastering. Any reliable engineer will take it from there, and apply the final limiting and EQ correction. Just always send along your limited version as well, so they know they have to do it at least as good as yours.

Finally, the only reliable situation in which we can actually use metering to check our mix is when using a master bus chain straight from the start. The spectrogram is only relevant when

the signal is peaking at zero. The loudness (or RMS) of various elements it's only relevant when the mix is peaking at zero.

Using the same, mostly “set it and forget it”, master bus chain

Many people like to work with an empty mix bus. It's fairly simple: every time the master meters peak over zero, they place an utility plugin on the master, and turn down the entire mix, and move forward. The problem I see with this kind of approach is that we can do this “floating” level thing forever. And we can't really anchor our mix. And then we end up at the mastering stage that we can't really push the mix because it falls apart due to unmanaged balance.

Now the reasoning above justifies the case for using a master bus chain, but **why make a “set it and forget it”**, when every song is different? Thanks for asking! While the songs are indeed unique, the ears that are listening are the same. Every time. We are not correcting things with the master bus chain, we are just stabilizing and enhancing the mix. Because the correction happens in the mix (since the chain is giving us the opportunity to hear every little detail), the chain will do a handful of stabilization tasks that are always the same. So there's not too much need for a vastly different chain for each song.

Alternatively, you can watch the processes (the multiband, the compressor, the limiters) and when you notice they're constantly triggered, or a band is working way too hard (which means it corrects), you can start investigating what track(s) are contributing to that overload, and start taking measures to mitigate the issue.

Building the master bus chain

By now you realized that, especially when we want to squeeze every little detail out of every track, starting with at least a limiter on the mix bus is important. With this method I have the opportunity to know from the early stages of the mix what's the level at which the kick sounds the way I like it to sound. If it gets buried in the mix, I know I can't make it any louder, because I hit the limiter's ceiling. I force myself to only work on turning down, panning, eq other elements, to make the mix work. **Having a clear delimitation for the canvas is a very precious piece of information.**

A good order of operations for building a master bus chain would be: harmonic additive effects, compression, imaging, equalization, (multi-band) limiting. The order of operations matters big time. For example, saturation can enrich a mix when done first in the chain, and absolutely mud it when done before the limiter.

Saturation → Compression → Imaging → Equalization → Limiting

| Saturation

I am not really a fan of characterful saturation for the master bus chain, because it tends to add too much “mush”. I get better results with saturators that are sparse in harmonics, like Inflator (it generates about 5-7 odd harmonics) or something like SPL MachineHead (this one generates just 1 extra harmonic). Especially the last one sounds like doing nothing, but it can shape the low end waveforms and deliver 0.7 to 1.2 dB of headroom. Clean headroom, with no dynamics artifacts and minimal harmonic distortion introduced.

Basically I don’t want to hear the saturation, but I want to gain more headroom.

Even a Fairchild compressor emulation (with the needle doing nothing - like Ozone Vintage Limiter) can be a good source of mastering saturation. Or transformer emulations. Or even a safety clipper, doing nothing for the vast majority of the time, but making sure the signal would not exceed 0 when things get hot.

Tape emulation is already debatable. Many times it can be too much for a good mix. Try it out, but be open to the possibility of removing it.

The key is to observe when the saturation is thickening the mix too much, or it reduces the crest factor (the difference between RMS and peak level in the low end) and dial it back. It can make the mix loud, but also can easily reduce definition, depth and contrast.

If in doubt, use no saturation. For sure you already using it on the tracks that need it.

A bit of “heads up” regarding tube saturation. Tubes will introduce even harmonics, which will impart asymmetry to the incoming waveform. This means that one half of the waveform (either the top either the bottom) will have its amplitude reduced. Applied as part of the master bus chain, tube saturation can actually decrease the loudness potential of the mix. Not all mixes need to be loud, but if you want a loud mix its better to replace tube saturation with a touch of soft clipping (this one will generate just odd harmonics, which are always perfectly symmetrical).

Black Box Analog Design HG-2 is always highly regarded. Or Shadow Hills compressor. Being tube based, they introduce asymmetry. Ozone Vintage Limiter (Fairchild tube emulation) will not. For loud music, this last one is better. Chose accordingly.

| Compression

I want to point out that treating each track individually (eq and compression) is very important, but it is not enough to solve all potential issues. When all tracks interact on the mix bus, they will create other side effects - new peaks, maybe frequency buildups - and these in turn will create new challenges for the limiter.

I don't see the bus compression as "the glue". The entire chain is "the glue". I use the bus compressor as a transient tamer for the occasional peaks. The previous stage (saturation) did a bit of rogue transients taming (but it can't do too much without audible distortion). Now the compressor will do the same but in a more transparent manner, since we took care of the very high rogue peaks.

At most, the entire mix should trigger 2 dBs of gain reduction on the compressor. I'd go with a 2:1 ratio and a threshold as close to 0 as possible - by "as possible" I mean avoiding distortion. I just want to tame the occasional peak, or to control the entire mix when a loud kick is hitting. I almost don't want to hear it work, but I want to see on the meters that I recovered a little bit of headroom.

Compression can also be used as a diagnostic tool: if it bounces too much, I'll find which tracks are too loud and act on them.

To better understand why small values matter, think about this way: if you want a loud song, and you need about 6 dB to sound as loud - and, more importantly, same audio quality (perceived dynamics, clarity, width) - as the reference, then realistically you have to have the limiter doing 6 dB of gain reduction. But that would be sonic suicide. A strategy is needed.

Every little move throughout the entire chain before the limiter - saturation, bus compression, multiband compression, clipping - that can transparently take away peaks and compact the sound, will result in the final limiter doing less and less. I usually end up with the final limiter doing 1 dB of gain reduction, by distributing the peak reduction and density throughout the components of the chain.

Imaging

Here I keep it simple. Another process to add a bit of something. Ozone Imager in 4-bands mode. The low band (under 100 Hz) does nothing. I always narrow the low end at the track level, not on the master bus. The low-mid band will have the biggest amount of width (in relation to the rest of the bands). Then the high-mids a little less width, then the highs even less.

Automating the width - narrower for verses and wider for choruses - can really add energy to a mix.

Equalization

EQ has a double role: to compensate for whatever changes in the tone the limiter may produce, and to make the final tweaks of the mix (usually high and low shelves, up to 1 dB in either direction). If you find yourself needing 3-4-5 EQ nodes, either go back to the mix, either let it go this time and try to make it better next time. We don't (have to) nail every mix perfectly.

I would also avoid doing M/S equalization trickery on the master channel (except for a bit of imager widening as previously mentioned). It's your mix, any unwanted residue (like too wide low end) can and should be corrected at the source. **We want every component of the master bus chain to only enhance the mix, not to fix it.**

Limiting

I like to use very fast settings for limiting (in case of Ozone Maximizer values between 0 and 0.2), and push the limiter for as long as distortion is not audible. Many limiters allow for "Delta" preview - to hear what elements are triggering the limiter the most. The next move would be to adjust their level, dynamics or EQ to mitigate their effect on the mix.

Of course if I need to squeeze one last dB, i'll be using dual limiters, a tiny amount of clipping or tools like Musik Hack Master Plan or Make Believe Studio CGII. Whatever sounds the cleanest for the mix at hand.

What's the input level into the chain?

Normally we'd want the plugins in the chain to work at their optimal range. I actually like to have a hot mix, close to 0 dB FS. Then the entire master bus chain will just round up peaks, compress, and make the entire mix denser.

In any case, I don't want the kick alone to trigger too much (if any) compression and limiting. Same for the vocals. Basically whatever played in solo, the master bus chain will hardly affect any element. Only when the entire mix is playing, that would be the moment when the dynamic processing kicks in. This can be achieved by carefully balancing the thresholds and mixing levels.

Here's a strategic move that always delivers. Read it as many times as you need. It makes a world of difference. After you gather some experience, you'll notice that the master bus chain adds a certain amount of loudness to a mix. Let's say 6 dB. If you know for a fact, that your chain will end up adding 6 dBs, you can feed it a darker mix. Pull back your hi-hats and other high-end-contributing elements with 6-9-12 dB. When this dark mix hits the saturation, it will not generate so many harmonics (because the high end is lower in level). So for the most part the master bus processing is dealing with rounding up the low end and low mids, gaining headroom. The high end will not be saturated or clipped, and you'll keep the mix sounding clear and with excellent transients. And when you hit the final limiter, the mix gets one final lift, and the high end will be placed at the desired level (6 dB louder), with minimal impact on the transients. If you smash a bright mix through the same set of processes (saturation, compression, limiting), you'll notice how undefined and "compressed" the result will sound.

People fail at building a proper (universal) master bus chain because they don't have a system, don't have a clear purpose for each component. It's not just about throwing plugins and hope for them to do their magic. This kind of chain is a big time saver. Work spending time understanding each kind of processing, and then make your own chain. And, can't stress this enough, don't rely on it for correction, it will ruin your mixes.

I am not exaggerating when I state that mixing is just gain-staging properly. When this fundamental aspect is done incorrectly, everything falls apart and we spend too much time on fixing the problems created by it. Always pay attention to the levels and suddenly every track and the mix will work properly with less plugins, and sometimes no plugins.

Good practice

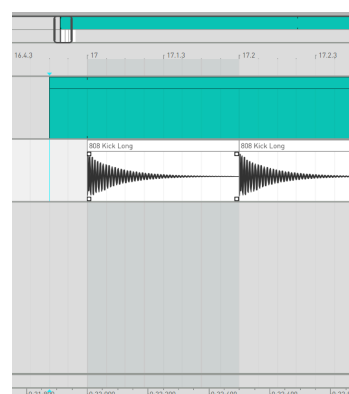
File management

First of all, almost every day someone announces they've lost everything because their drive failed, computer broke, or got stolen. **KEEP BACKUPS!** I have a cloud backup for projects, and I do a local weekly backup for everything (including projects). It's very unlikely for the backup drive and system drives to fail all at the same time. There are many options from portable drives to NAS units, make sure you have at least one backup solution installed.

Next, get organized with your projects. I keep a simple structure by year, and then the projects themselves start with a 4-digit number and then artist name - track name (version). If I work on a project and I have no idea for the name, the unique 4 digits are enough to keep everything in order. No more projects with random names everywhere on the hard drive.

Exporting the project - pre-roll and fade-out

When an audio file is manipulated (including encodings), it may lose few ticks from the beginning and or the end. Often times, people mention that they somehow lost small parts of the first transient of the first kick. For this reason, it's important to leave a bit of empty space (20-50 ms) at the beginning of the mix. Most DAWs allow for adding "negative bars" at the beginning of the project. The project will still start at bar 1 (so no LFO synced to the beat are lost), but when exported, we get that extra safe space. Ableton Live doesn't allow for this, so the best way to work around it is to start the project at bar 4, 8 or 16, and when exporting, to set the In marker slightly before the first beat. Also leave enough space for the tail of the outro fade out completely.



Allocating a short blank pre-roll area in Ableton Live

What I like to do, is to use an empty midi track as guide. I draw a long midi region to include the extra space at the beginning, and the tail at the end. I also automate the master fader to fade out the volume at the end of the song, to make sure the end of my midi guide is with the fader completely down. And when exporting, I simply select the guide, and press the key command to loop it, and I instantly have the selection (including pre-roll and fade out) ready for export. I usually do the radio version and extended version in the same project, so I also label the guide region with the version it represents.

Sample rate

The choice of sample rate doesn't really matter. Some sound cards work better (less clicks and pops) at 48 kHz, while others at 44.1 kHz. Pick whatever works best, and record and export at this sample rate.

In case you need to convert to a different sample rate, use a specialized tool like [r8brain](#) to upsample or downsample audio. Preferably use a source at 32-bit floating point no dither.

Bit depth

The most compatible by far, and often the only one accepted by music distributors is 16-bit. When exporting in 16-bit, that means the file is the final master. Apply dither when exporting. Then you can convert it to FLAC, MP3, M4A etc.

24-bit has slightly more dynamic range than 16-bit, but can already create compatibility issues with some playback systems. Mainly used for distribution for what's called "HD Audio".

32-bit floating point never clips the audio regardless of how much the level crosses over 0 dB FS. It's the safest format to send files for mastering, or for further processing before being exported in the usual 16-bit.

Dither

Dither is just (white) noise with a very low amplitude. Its level depends on the bit depth, and it's always at the lowest level that can be represented by that particular bit depth ("noise floor"). For 16-bit audio, the dither level is at around -96 dB.

We apply dither noise only when we do bit depth conversion from a higher bit depth to a lower one. 32-bit floating point due to its nature, doesn't require dithering. 16-bit and 24-bit they both do. When dealing with 24-bit audio, dithering multiple times (if the file is edited multiple times) is not an issue. However, when dealing with 16-bit audio, it's good practice to apply dither just once, during the final export.

A good workflow can be exporting in 32-bit floating point, master this file, and then export to 16 and/or 24-bit with dither.

In the vast majority of cases, using or not using dither makes no audible difference. If you want to put this topic to rest once and for all, I suggest "[WTF is Dither?](#)" and "[Only Dither Once](#)" by Dan Worrall (YouTube videos).

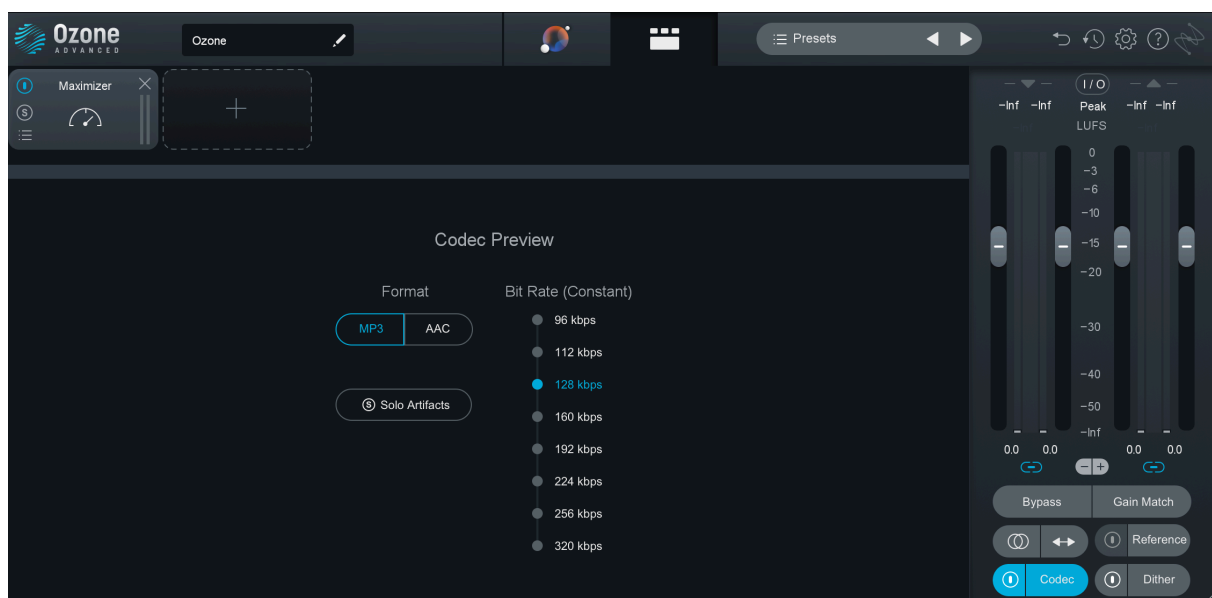
Audio file formats

We have lossless and lossy file formats. Always keep a lossless copy of your finished music.

Lossless will preserve every bit of original audio: WAV, FLAC (Free Lossless Audio Codec), ALAC (Apple Lossless Audio Codec), AIFF (Audio Interchange File Format).

Lossy will discard unimportant bits of audio (masked signals, or very high frequency Side information) in order to preserve as much audio as possible. MP3 (specifically the LAME encoder at 320 kbps ABR Joint Stereo) is still the best codec out there. The new generation of AAC codecs behave better than MP3 when talking about lower bitrates. But when it comes to strictly 320 kbps, the MP3 is the best bet.

If you're looking for a complete audio encoding and conversion solution, [fre:ac](#) is one of them.



iZotope Ozone 10 Mothership

In iZotope Ozone's mothership plugin you can use the Codec section to preview your master through various audio codecs at various quality settings. Or you can use it as a lo-fi effect, by using a low bit rate setting.

The end

The end.