Synthesized Bass

By Bohdan Raczynski

So, when played on the new hi-fi system, your favourite music sounds overly bright ?. Perhaps it is due to poor recording practices, or perhaps the bass player himself did not explore his instrument fully?. This seems to be generally true, even for some great artists recorded in 60's and 70's. Contemporary recordings appear definitely better balanced tonally, but even so, the device described in this article may add unexpected dimension to your listening pleasure.

It would be desirable to be able to fully exploit the benefits of good quality loudspeaker system. A well-designed system of today, can deliver broad-spectrum sound, easily reaching down to 20Hz. But how often do you actually hear a really deep bass note coming out of your subwoofer?. In fact, the loudspeaker system can well expose all shortcomings of the recording studio practices, be it too much undesirable noise, or poor tonal balance.

Graphic equalizer is perhaps a helpful tool in correcting some of the issues referred to above. Even so, I found myself re-adjusting the knobs from one song to another. Settings, that worked for one song, seemed to be doing the opposite for another.

So, is there anything at all, that could help in this situation?. Perhaps the answer is provided by the music itself.

A little digression into music

Musical **octave** is the interval between one musical pitch and another with half or double its frequency. For instance, if one note has a frequency of 100Hz, the note an octave above it will have 200Hz frequency, and the note an octave below will have 50Hz frequency. Interestingly, the human ear tends to hear both notes as being essentially the same, due to the same harmonic relationship. Not surprisingly, notes an octave apart are given the same note name. This is called octave equivalency, and assumes, that notes one or more octave apart are musically equivalent.

Given the above – what if we could make use of the low-frequency notes as a guiding reference for synthesizing the same notes, but one octave lower ?.

Sub-harmonics addition - this is the basic idea, behind enhancing the low-frequency content of the existing recordings.

Back to technical issue

Implementing FIR filtering techniques on a personal computer with a sound card is so well documented in the available literature, that I will only briefly mention it's existence later.

Standard pitch-shifting algorithm (also known as Phase Vocoder) is based on a two-stage processing. In Stage 1, or analysis stage, a sound sample is windowed and then a short-time Fourier transform (STFT) is applied to each segment, or "frame". Typical values for the length of the frame are: 1024, 2048, 4096 and 8192 samples or "bins". The resulting spectrum can be manipulated accordingly to the expected results and then re-synthesis, or Stage 2, is accomplished by applying the inverse STFT to each segment. This will return the sound sample back to time domain, and the segment is ready to be played out by the sound system. In order to avoid discontinuities at the frame boundaries, the segments/frames need to be windowed after re-synthesis and overlapped, which means, that much of the processed signal is fed back to the input and re-processed again and again, depending on "overlap factor". The overlap factor tells you how many times the frame is re-processed. This process is known as WOLA – Weighted OverLap Add. Both: analysis and synthesis stage typically attempt to use amplitude and phase information in each of the FFT bins. It should be clear by now, that analysis of the frame and synthesis of the sample are done in frequency domain.

Unfortunately, low-frequency sound quality produced by such method is questionable. Sound notes parameters such as "attack", "sustain" and "decay" all suffer degradation. The "attack" is slow and the sound builds itself over several frames. The "sustain" suffers from amplitude fluctuations and exhibits a fair amount of echo. The "decay" is too slow and smeared.

Overall, the sound suffers from "phasiness", echo, and amplitude instabilities. Much of the amplitude instabilities are attributed to lack of phase synchronization between FFT bins. That is, if the frequency domain analysis stage uncovered a bass note that falls on the boundary of two bins, the phases of those bins are not synchronized, as the bins know nothing about each other.

I have experimented with standard PV algorithm and tried various possible improvements, but finally discarded it as unsuitable for the bass synthesiser.

Somewhat different approach

There are two processing streams within the bass synthesizer. Standard sound processor has 5 processing units: selection of high-pass filters, two shelving filter selections, a selection of low-pass filters and finally Q-notch/peak filter. Each filter has several characteristic parameters to choose from. This array of tools allows for much more professional sound processing than bass/treble set of controls on a typical amplifier, or even graphic equalizer. One can cut excessive sub-audio, low-frequency noise without actually cutting off useful bass information. A 50Hz hum can be removed with Q-notch filter. Furthermore, one may prefer to reduce "harshness" or add "brightness" to the favourite recording.

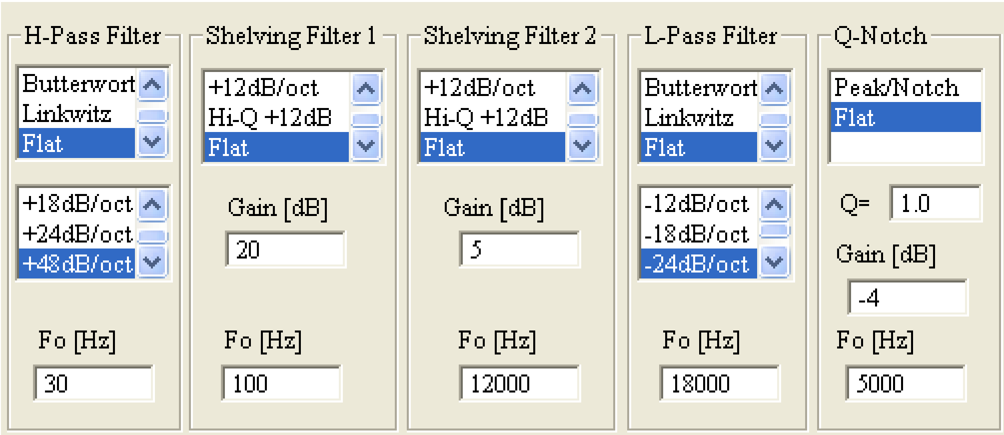

Figure 1. DSP filtering section

Bass Synthesizer

The second stream is the bass synthesizer. Algorithm implemented in the bass synthesizer is also a two-stage process. Collected sound sample is FFT transformed into frequency domain for analysis. The algorithm only scans frequency range between two nominated limits: F1 and F2.

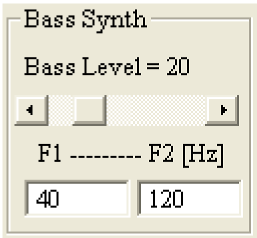

Figure 2. Very simple bass controls of the bass synthesizer.

However, the analysis stage discards information about the phase of each FFT bin. The technique used here is known as STFTM – Short Time Fourier Transform Magnitude. The length of the frame is selectable as: 4096 or 8192 samples. The synthesis stage is however performed in the time domain. The synthesized waveform is then transformed into frequency domain, where is low-pass filtered and then mixed into both left- and right-channel sound streams. The synthesized waveform is always started with phase=0 degrees, so the “attack” time is practically zero. All the above happens within single “frame”. The phase of the synthesized waveform is maintained from one frame to another. The overall processing algorithm uses standard “Overlap-and-Add” DSP technique, but it does not employ the WOLA approach. Since there is no re-processing of synthesized frames, the resulting sound is free from “echo” and “phasiness”.

Operation of the bass synthesiser in frequency domain is shown on several spectrograms below. Figure 3 depicts a sample spectrogram of a well-known pop-music artist. The song comes from “A hard Day’s night” CD, and is titled: “I should have known better”, The Beatles, 1964 EMI records LTD. It is immediately observable, that low frequency content of this recording is very poor. The song itself is great, but to the listener’s ear, the music sounds too bright and ill-balanced.

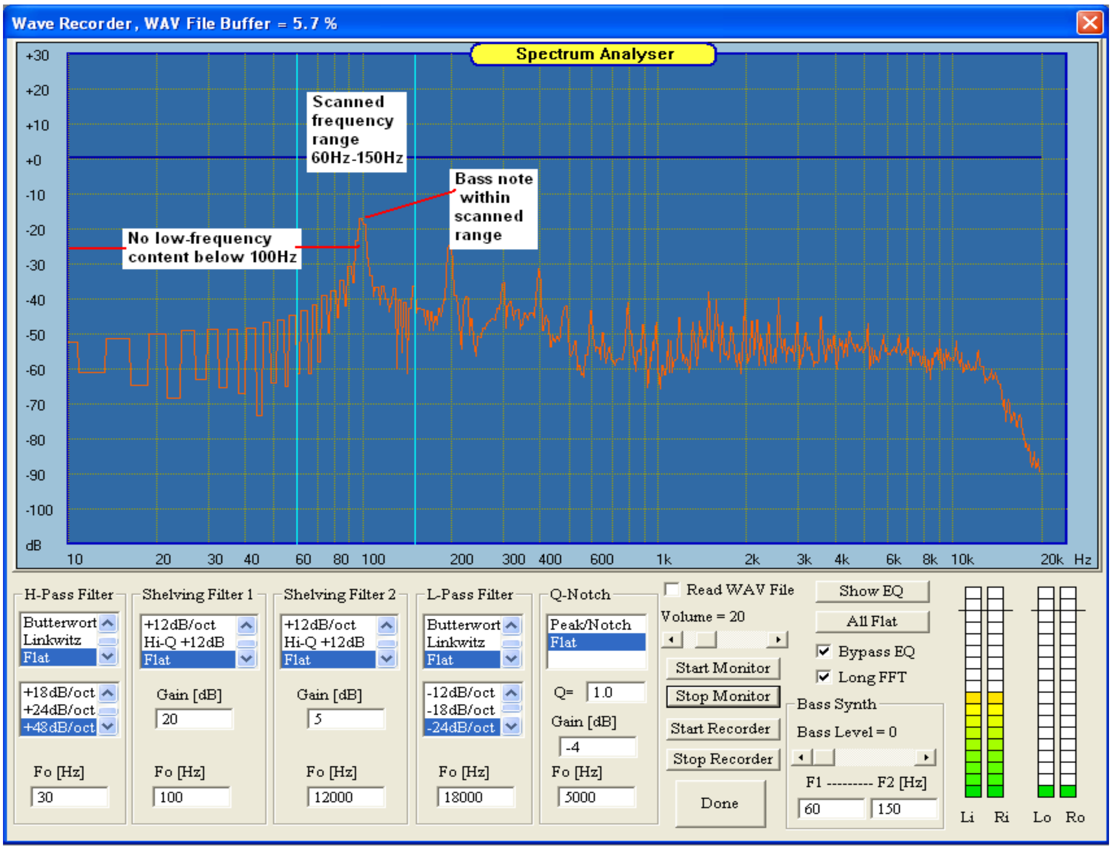

Figure 3. Sample spectrogram of "I should have known better" – The Beatles.

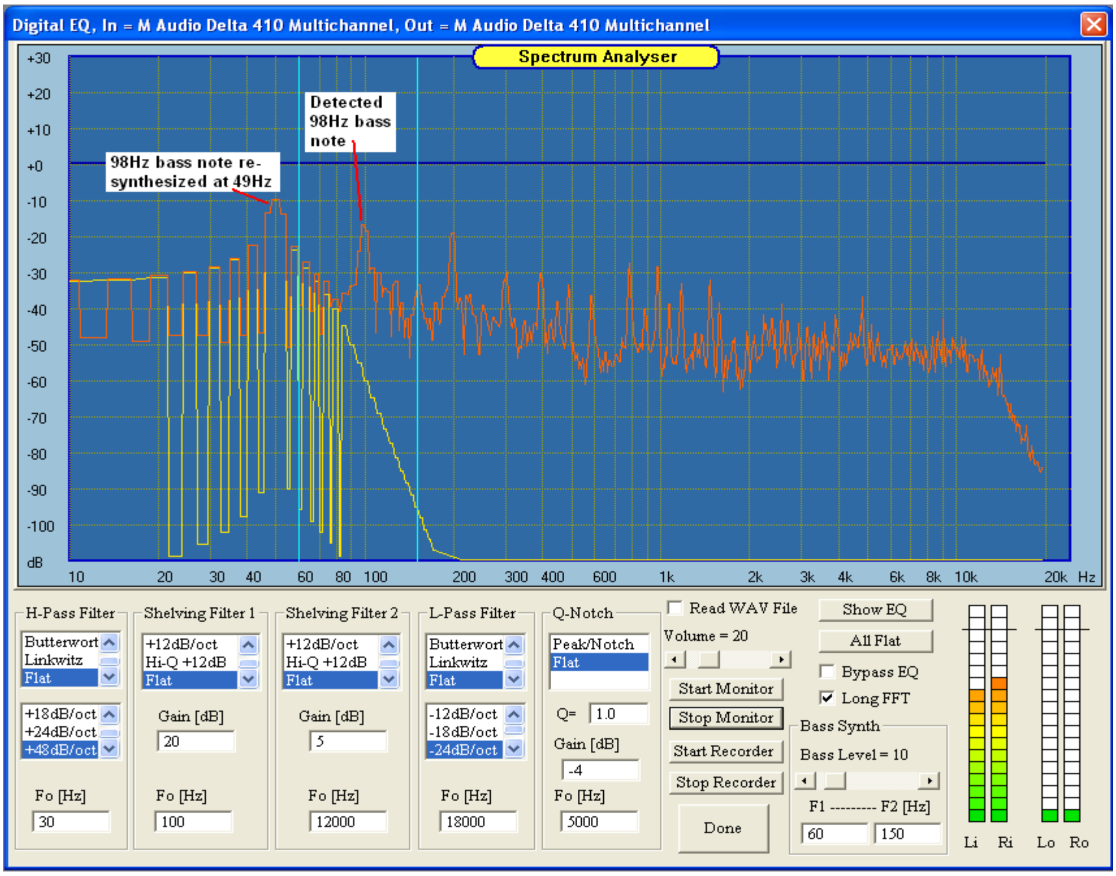

Figure 4. Same song as on Figure 1, but with bass synthesizer turned ON.

The bass synthesized can work with already very low frequencies. A good example is shown below, where a 36Hz bass note is re-synthesized down to 18Hz. At this point it would be important to realize, that the bass synthesizer can inflict serious damage to subwoofer. For instance, by setting low frequency boundary to 20Hz, you allow the synthesizer to generate 10Hz sub-harmonics at high amplitude.

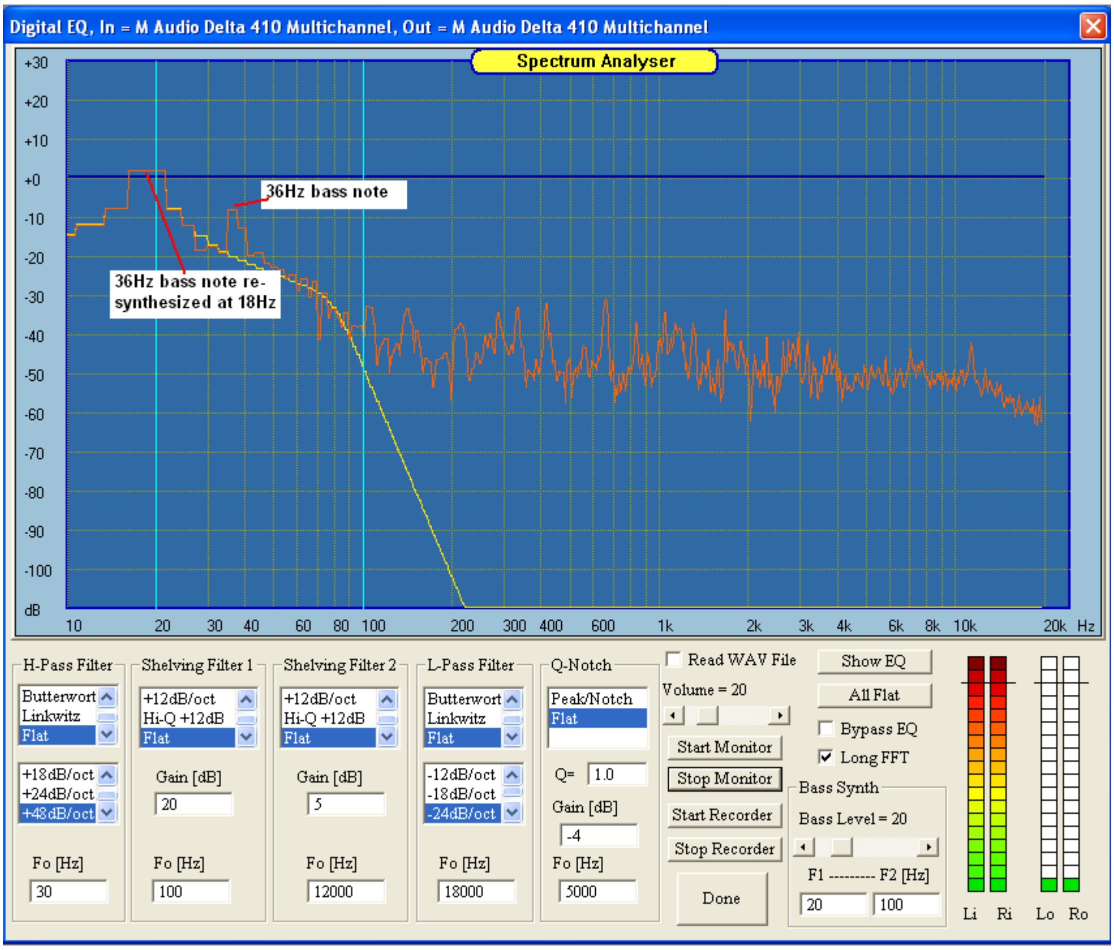

Figure 5. Artist: George Michael, song called “Spinning the wheel”, comes from CD “Older”.

How does it actually sound

Good. The reason for it is, that the only thing that is different from the original recording is the addition of a sub-harmonic to the lowest bass notes in the recording. The vital amplitude-phase relationship of the original harmonics in the bass note has not only been preserved, but also enhanced with an extra sub-harmonic, one octave down.

In order to visualize the above, I have temporarily modified the program to allow the synthesized sub-harmonics to be outputted from the sound card on a separate channel. Therefore, the music channel and the sub-harmonic channel can be simultaneously captured on a digital storage oscilloscope for comparison – as shown on Figure 6.

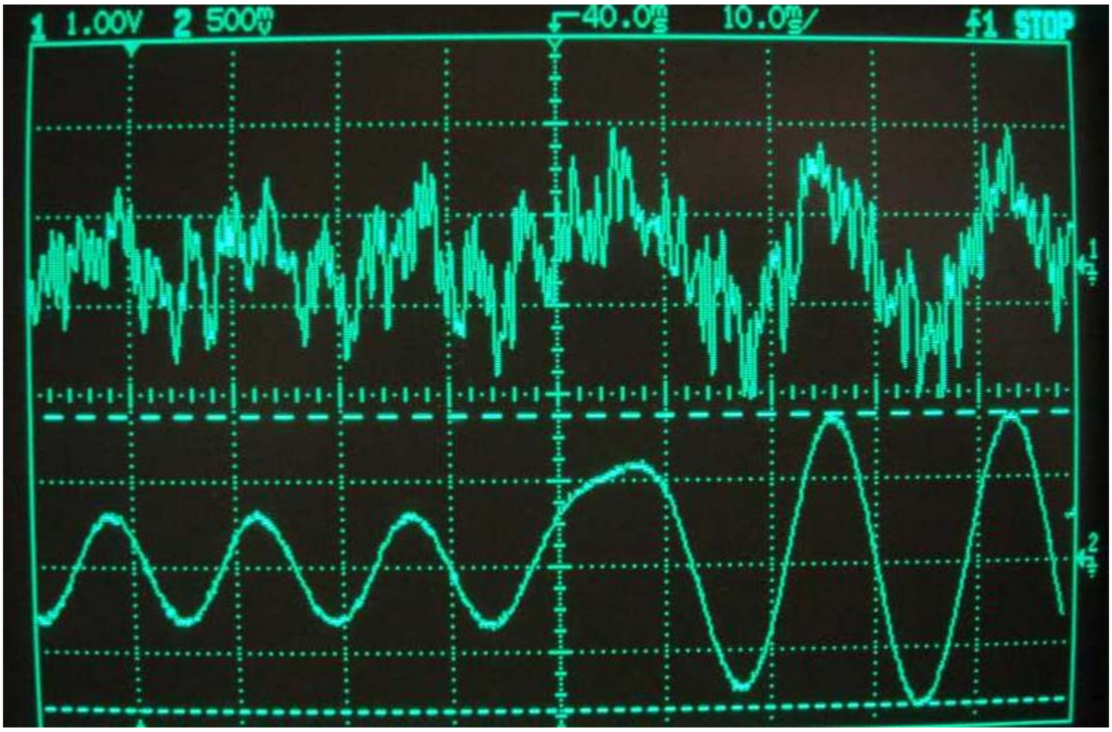

Figure 6. Trace 1 – music channel + sub-harmonic channel, 1V/dev
Trace 2 – sub-harmonic channel alone, 0.5V/dev.

It can be observed on Figure 6, that the sub-harmonic, generated as a pure sine-wave, takes on characteristics of a live, instrument-generated sound, when combined together with the original set of harmonics. This is evident in the fact, that the sine-wave has changed it's smooth, "mathematical" shape and after adding the original set of harmonics, it now fully contributes to the original instrument's sound.

Additional and desired characteristics relate to the envelope of the synthesized sub-harmonic, namely: zero attack time, fast amplitude updates every 92ms (this is due to the STFTM frame length, which is 4096 bins and sampling frequency of 44100Hz so that frame length in time domain is: 4096/44100 =92ms), and instantaneous decay. All these characteristics contribute to maintaining the "musical" character of the final sound.

Interestingly, Figure 6 captures transition between two different frequencies and different amplitudes of bass notes. The uninterrupted transition within 10 ms is clearly evident, indicating high level of performance of the synthesizer.

Conclusions

The device described in this article incorporates several audio DSP filters, a bass synthesizer and can record into a WAV file. Thus allowing the user to create his own, perfectly balanced, bass-enhanced ".wav" file recordings, that can be made into a standard CD using any off-the-shelf CD recording software, like "RecordNow".

The bass synthesizer works best if the musical content has single, well defined bass notes within user selected frequency range. The resulting new bass notes actually sound quite musical, despite being completely synthetic. Perhaps mixing them back into the original sound stream makes them feel like the real artist playing. DSP processing artefacts are not audible during typical processing application.

By now, I have re-recorded all The Beatles song catalogue for my own use. I have actually stopped listening to the original CDs, as they sound too poor in comparison with the restored and enhanced material.

So, are you ready to keep your subwoofers really busy?.