Sound Comparison Analyzer ver.4.1.0, User Manual

Hwang Il Seok[1]

[1] *H.I.S. Violin Atelier* / www.hisviolins.com / hisviolins@gmail.com

November 21, 2026

This program shows the difference between the frequency characteristics of two sounds as a graph. If you take the spectrum (level at each frequency) of each sound and subtract one from the other, you get the difference at every frequency. However, the spectrum of a sound is very complex, so the difference of two spectra is just as complex and hard to interpret.

Therefore the program first finds the envelope of each spectrum and then subtracts the envelopes. An envelope is a smooth curve that shows the overall shape of the spectrum, so their difference is also one simple line, and you can easily read at which frequencies the sound became louder or quieter, and by how much. This is the same as a difference spectrum, or the transfer function in dB with X as the output and Y as the input.

The calculation is as follows.

1. Find the frequency spectrum (FFT : Fast Fourier Transform) of X and of Y.
2. Find the envelope of each spectrum.
3. Subtract the envelope of Y from the envelope of X to find the difference (X − Y).

The envelope is found with one of LPC, True Envelope, Smoothing, MGCE and LTAS[1] (section 5.2). The result graph is read by dividing it into amplified ranges (Live Zone), attenuated ranges (Dead Zone) and so on, and the program finds the Zones and Spots automatically and shows them in the graph and in a table (sections 5.6 and 7).

[1] LPC (Linear Predictive Coding) : models the sound as a resonant filter; mainly used in speech analysis. True Envelope : an envelope through the tops of the harmonic peaks. Smoothing : averaging over frequency bands. MGCE (Mel-Generalized Cepstrum Envelope) : an envelope found on a frequency axis stretched in the bass like the human ear. LTAS (Long-Term Average Spectrum) : the average spectrum of the whole file. See section 5.2 for details.

1 Usage

Use it when you want to know the difference between two sounds or how a sound has changed. Examples for the violin :

- Change of sound with the position of the soundpost
- Change of sound with the shape of the bridge
- Change of sound with new strings
- Acoustic characteristics of the violin body alone (the same string recorded with and without the body, and compared)

Other ways to measure the acoustic characteristics of the body include the tap tone (main resonances) and bridge admittance measurement (with special equipment such as an impulse hammer). This program compares two sounds recorded with one microphone, so it needs no special equipment and can also be used for other instruments and for sounds other than instruments.

2 Installation

This is a portable version, so no installation is needed. Unzip the file and double-click
"Sound_Comparison_Analyzer_v.X.X.X.exe" in the folder
"Sound_Comparison_Analyzer_v.X.X.X". Depending on the computer, starting may take some time.

3 Preparations

3.1 Preparation of the sound files

Prepare the sound files to analyze. Up to 4 pairs (X1-Y1, X2-Y2, X3-Y3, X4-Y4) can be loaded, and the result is always shown as X − Y. Therefore **put the sound to be compared (the later sound, Target / After) in X and the reference (the earlier sound, Reference / Before) in Y**.

- Change of the soundpost position : X = sound after moving it, Y = sound at the reference position
- Amplification by the violin body : X = sound of the violin with the string mounted, Y = sound of the string itself without a body

To compare the tone of the whole instrument, analyze a sound of many notes (a scale, part of a piece, etc.) with the LTAS method (section 5.2).

3.2 Conditions of the sound files

The sound files must meet the following conditions.

- Exact tuning and exact pitch
- Uncompressed wave file (extension : wav). 8, 16, 24 and 32 bit integer and 32 bit floating point formats are all possible
- Mono (1 channel) recommended. A stereo file is converted to mono by averaging its two channels
- Sampling frequency of 44.1 [kHz] recommended. All files calculated together must have the same sampling frequency
- Longer than the Analysis Length (default 1.0 s); 1 s or more recommended

3.3 Precautions

When recording, keep the resonance of the room out of the recording. An anechoic room is best; otherwise a quiet outdoor place (a field with no buildings that reflect sound, etc.) or a room well damped with curtains is recommended. If the room resonance has been recorded with the sound, take it into account in the analysis.

Always record X and Y with the same microphone, at the same position, distance and recording level. With different microphones the difference of their characteristics remains in X − Y, and with different recording levels the whole result graph moves up or down and the Zone judgement goes wrong.

4 Screen layout

When the program starts, the screen of Figure 1 appears. The left column holds the sound files, the envelope method and the FFT settings; the right column holds the X-axis, the mic calibration, the Zone / Spot detection and the calculation buttons. The numbers of the items show the rough order of use.

No.	Item	Contents
1, 2	SELECT SOUND SET X / Y	Sound files to compare (4 for X, 4 for Y)
3	ENVELOPE METHOD	Method of finding the envelope (5 methods), its options, a short description
4	FFT SETTINGS	FFT Size, Window, Analysis Length
5	RANGE OF X-AXIS	Scale (Linear/Log) and range of the X-axis of the graphs
6	MIC CALIBRATION	Microphone calibration
7	ZONE / SPOT DETECTION	Automatic judgement of Zones and Spots (display, T, Q, Spot height)
8	CALCULATION	Graph window mode, calculation buttons

Table 1: *Screen layout*

The screen has short descriptions; please refer to this manual for details. The color of a button shows its state.

Color	Meaning
Sound X1	Input needed (no file loaded yet)
Sound X1	Loaded / ready to calculate
LTAS + Smoothing	Selected envelope method, and a button used only with that method (X all − Y all)
X1 - Y1	Not available now (or an envelope method not selected)

Table 2: *Meaning of the button colors*

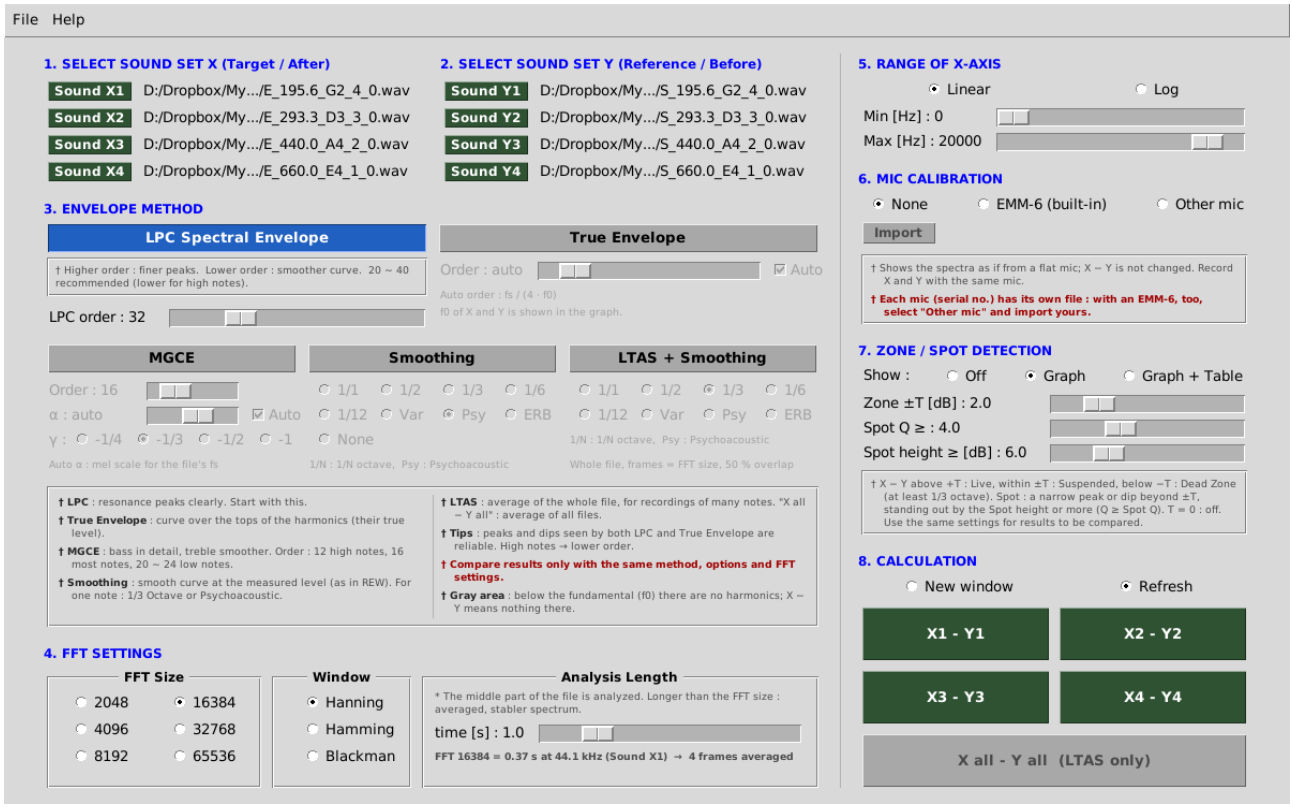


Figure 1: *Program screen (8 sound files loaded)*

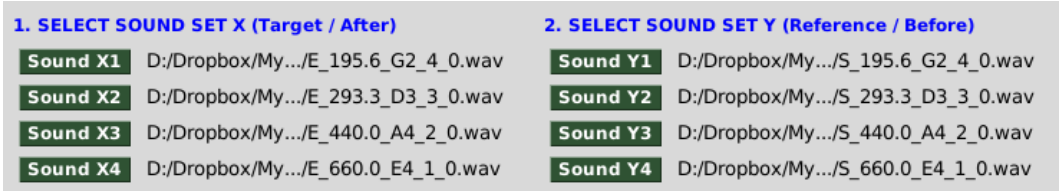
5 How to use

5.1 1, 2. SELECT SOUND SET X / Y

Click [Sound X1] to load a sound file, and load the other files in the same way. When a file is loaded, the button turns from orange to green and the path and name of the file are shown on the right (Figure 2). The calculation is done in the pairs [X1 − Y1] ∼ [X4 − Y4], so load the files accordingly.



(a) *Before loading*



(b) *After loading*

Figure 2: *Loading the sound files*

A message box appears in the following cases.

- Stereo file : the file is converted to mono for the analysis. Warning that the level is about 6 [dB] lower if one channel is empty
- File shorter than the Analysis Length : the whole file is analyzed (not shown with LTAS)
- Unreadable file : an error message appears and the file loaded before is kept

5.2 3. ENVELOPE METHOD

Selects the method of finding the envelope of the power spectrum. When a method button is pressed, only the options of that method become active, and the X-axis of the graphs changes to the scale (Linear / Log) that suits the method.

■ **Resolution and the X-axis :** resolution means how narrow a feature (in Hz) the curve can follow. LPC and True Envelope have an even resolution in the bass and the treble, so they are shown on a Linear axis. MGCE, Smoothing and LTAS are fine in the bass and coarse in the treble, so they are shown on a Log axis. A Log axis spreads out the bass and compresses the treble, so on the screen these curves also look evenly detailed.

Method	Principle	X-axis	Bass res.	Treble res.	Suited for
LPC	Resonant filter model	Linear	Even	Even	Resonance peaks
True Envelope	Curve over harmonic tops	Linear	Even	Even	Harmonic levels
MGCE	Bass-stretched axis	Log	Fine	Coarse	Bass resonances
Smoothing	Band average	Log	Fine	Coarse	Measured-level trend
LTAS	Whole-file average + Smoothing	Log	Fine	Coarse	Many notes, many files

Table 3: *Methods of finding the envelope (res. : resolution in Hz)*

■ LPC (Linear Predictive Coding)

- Principle : regards the sound as a resonant filter and draws the response of that filter. The curve follows the resonance peaks and is little affected by noise or by the valleys between the harmonics.
- Resolution : even in the bass and the treble.
- Settings : [LPC order] (1 $\sim$ 128, default 32). The higher the value, the narrower the peaks it follows; too high, and the envelope becomes almost the power spectrum itself. 20 $\sim$ 40 is recommended; for high notes it starts to follow single harmonics from about 48.
- Suited for : seeing the position of resonance peaks clearly. Dips are less accurate.

■ True Envelope

- Principle : repeats cepstral (see MGCE) smoothing to find a curve through the tops of the harmonic peaks. It is not pulled down by the valleys between the harmonics.
- Resolution : even in the bass and the treble, and coarser than the harmonic spacing (about 4 times the harmonic spacing). Narrower resonances do not appear.
- Settings : with [Auto] (default) the program finds the fundamental (f0) and sets the order to fs / (4 $\times$ f0) (fs : sampling frequency). The order is kept low to suit the harmonic spacing (f0), so that the curve connects the tops of the harmonics instead of following each one. For example, in a 44.1 [kHz] file G3 (196 [Hz]) gives 56 and A4 (440 [Hz]) gives 25. The f0 found is shown in the graph. For sounds without a confirmed harmonic structure (tap tones, etc.) no f0 can be set and a message box appears; then turn

Auto off and set the [Order] (10 ∼ 200) by hand.

- Suited for : the actual level of the harmonics (the level of their tops)

■ MGCE (Mel-Generalized Cepstrum Envelope)

- Principle : draws the envelope from the first few coefficients of the cepstrum (the Fourier transform of the logarithm of the spectrum). It is calculated on a frequency axis adapted to the hearing scale (Mel scale), so the bass is finer.
- Resolution : fine in the bass, coarse in the treble. At the same order, in a 44.1 [kHz] file it is about 3 times finer than LPC below about 2 [kHz] and coarser than LPC above about 7 [kHz].
- Settings : [Order] (4 ∼ 64, default 16) : about 12 for high notes, 16 for most notes and 20 ∼ 24 for low notes are recommended; too high, and the curve follows single harmonics. [α] is how much the frequency axis is stretched in the bass and compressed in the treble; with [Auto] (default) it is calculated for the sampling frequency so that the axis comes closest to the hearing scale (Mel scale) (0.544 at 44.1 [kHz]). [γ] is the type of analysis (default $-1/3$) : $\gamma = -1$ with $\alpha = 0$ is the same as LPC, and the closer to 0, the smoother the curve.
- Suited for : seeing the bass resonances in detail

■ Smoothing

- Principle : averages the level of the band around each frequency. The curve lies at the actual measured level and follows all peaks and valleys wider than the band.
- Resolution : fine in the bass, coarse in the treble; how much depends on the option (Table 4).
- Settings : choose from 1/1 ∼ 1/12 (1/N Octave), Var (Variable), Psy (Psychoacoustic, default), ERB (Equivalent Rectangular Bandwidth) and None.
- Suited for : the rough trend at the actual measured level

Option	at 100 [Hz]	at 1 [kHz]	at 10 [kHz]	Note
1/3 Octave	23 [Hz]	230 [Hz]	2.3 [kHz]	1/1 ∼ 1/12 in the same proportion
Var	1.4 [Hz]	120 [Hz]	2.3 [kHz]	Finest in the bass
Psy (default)	23 [Hz]	120 [Hz]	1.2 [kHz]	Peaks weighted
ERB	35 [Hz]	130 [Hz]	1.1 [kHz]	Bandwidth resolved by the ear. Plain average
None	FFT result as it is			

Table 4: *Width of the band averaged by each Smoothing option*

■ LTAS (Long-Term Average Spectrum)

- Principle : divides the whole file into frames of the FFT size (50% overlap), averages them, and then applies the selected Smoothing once. With a sound file of many notes the gaps between the harmonics are filled and the body resonances appear as a continuous curve. With a sound file of one repeated note the gaps are not filled.
- Resolution : follows the selected Smoothing option (default 1/3 Octave).
- Settings : the whole file is used, so the Analysis Length is inactive. The [X all − Y all] button averages several files; the LTAS of each file is found first and the files are averaged with equal weight, so the length of a file does not affect the result.
- Suited for : sounds of many notes, averages of several files

■ **Defaults of Smoothing and LTAS :** Smoothing is mainly used for a sound of one note and LTAS for a sound of many notes, so their defaults differ. The spectrum of one note is empty between the harmonics, so a plain average such as 1/3 Octave also averages these gaps in the treble, where the harmonics are far apart, and the curve tends to sag. Psychoacoustic weights the peaks and sags less, so it is the default of Smoothing. In LTAS the gaps between the harmonics are already filled and the peaks need no emphasis, so 1/3 Octave, which averages the energy as it is, is the default. 1/3 Octave is a resolution commonly used in acoustic measurement and equals the minimum width judged as a Zone.

■ How to choose a method

1. First choose by the kind of sound.

- Sound of many notes (a scale, part of a piece, etc.) → LTAS
- Sound of one sustained note → LPC, True Envelope, MGCE or Smoothing (see 2 below). LTAS can also be used; it is then the same as Smoothing of the whole file instead of the Analysis Length.
- To view several files together (even files of one note) → [X all − Y all] of LTAS

2. For a sound of one note, choose by what you want to see.

What you want to see	Method	Feature
Resonance peaks, clearly	LPC	Sharp peaks. Valleys less accurate
Actual level of the harmonics	True Envelope	Not pulled down by the valleys
Bass resonances in detail	MGCE	Bass 3× finer than LPC (same order)
Rough trend at the measured level	Smoothing	Smooth curve at the actual measured level

Table 5: *Choosing a method for a sound of one note*

3. Points to note

- First look with LPC (default) and check with True Envelope. Peaks and dips that appear in both are reliable; those seen in only one may come from the method itself.
- The spectrum of one note has information only at the frequencies of its harmonics. The higher the note (E string, etc.), the wider the harmonic spacing and the less information, so a lower order than the default is better. With too high an order the envelope follows single harmonics and becomes jagged.
- For Smoothing, 1/3 Octave or Psychoacoustic is recommended. Finer options fall between the harmonics

and make the curve jagged, so they suit only an LTAS of many notes. Choose Psychoacoustic to emphasize the peaks, and 1/3 Octave or ERB to see the actual energy without emphasis.

- Results to be compared must always be calculated with the same method and options (order, FFT Size, Analysis Length, etc.). With a different method or options, even the same sound gives a different result.

■ **Gray area :** for a sound with harmonics, the part below the fundamental (f0) is shown in gray (section 7).

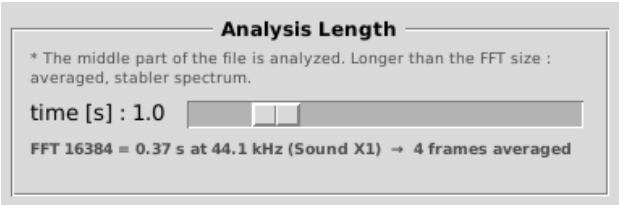
5.3 4. FFT SETTINGS

■ **FFT Size :** FFT (Fast Fourier Transform) is the frequency analysis, and the FFT Size is the number of samples analyzed at once. The larger it is, the higher the frequency resolution (when the analyzed part is at least the FFT Size). At 44.1 [kHz], 16384 (default) is 0.37 s and 65536 is 1.49 s.

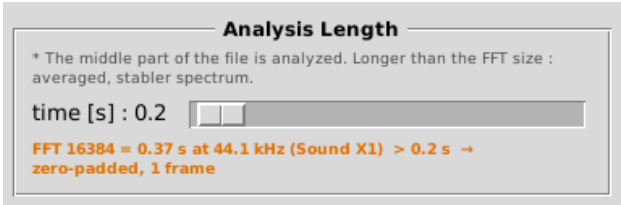
■ **Window :** the shape with which the data to analyze is cut out. It makes very little difference in the result; Hanning (default) is usually used.

■ **Analysis Length :** only this length (0.1 ∼ 5.0 s, default 1.0 s) from the middle of the sound file is analyzed. If it is longer than the FFT Size, frames of the FFT size (50% overlap) are averaged, so the spectrum is stable and has fewer noise peaks. If it is shorter, it is filled with zeros up to the FFT Size and analyzed as one frame; this only adds points to the spectrum and neither separates close peaks better nor reduces noise. If the file is shorter than this length, the whole file is analyzed.

The status line under the slider shows the time of the FFT Size and the type of analysis, calculated with the sampling frequency of the first loaded file (with no file, 44.1 [kHz] is assumed and (assumed) is shown). When frames are averaged, their number is shown in gray; when zeros are filled in, zero-padded is shown in orange (Figure 3). If it is orange, reduce the FFT Size or increase the Analysis Length.



(a) *FFT 16384 : average of 4 frames*



(b) *Analysis Length 0.2 s : zero-padded*

Figure 3: *Analysis Length and the status line*

5.4 5. RANGE OF X-AXIS

Sets the scale ([Linear] / [Log]) and the displayed range ([Min], [Max], 0 ∼ 22,050 [Hz]) of the X-axis of the graphs. The default range is 0 ∼ 20,000 [Hz]; on a Log scale, if Min is 0, the axis starts at 20 [Hz]. The Y-axis is automatic.

The scale only changes how the result is shown, not the analysis. When an envelope method is selected, the scale changes automatically to the one that suits the resolution of that method (Table 3); it can be changed by hand afterwards.

- Linear : LPC, True Envelope. Even resolution in the bass and the treble. Changed to a Log scale, the bass is stretched and the curve looks coarse.

- Log : MGCE, Smoothing, LTAS. Fine resolution in the bass, coarse in the treble. Changed to a Linear scale, the bass looks jagged and the treble looks smeared.

To see low-frequency resonances in detail, a Log scale also helps with LPC. The automatic judgement of Zones and Spots is done only within the Min ∼ Max range of the X-axis (section 5.6).

5.5 6. MIC CALIBRATION

Every microphone has its own frequency characteristic. With a calibration file, the spectra and envelopes are shown as if measured with a microphone with a flat characteristic. What is corrected is the deviation at each frequency (the shape of the curve), not the absolute sound pressure level.

- [None] (default) : no correction
- [EMM-6 (built-in)] : built-in calibration of the Dayton Audio EMM-6 (the values of the author's microphone)
- [Other mic] : load the calibration file of another microphone with [Import]. With no file the button is orange; once loaded, green (Figure 4).

[**Important**] Calibration values differ from microphone to microphone (from serial number to serial number), even for the same EMM-6 model. The built-in values are those of the author's microphone, so **even with an EMM-6, always select [Other mic] and [Import] the calibration file for the serial number of your own microphone.** The calibration file can be downloaded from the manufacturer's website by entering the serial number.



Figure 4: *Other mic selected (before importing a calibration file)*

The calibration file is a text file (.txt, .cal, .frd) with two columns, [frequency (Hz) correction (dB)]. Lines beginning with *, ", #, ; and lines that are not numbers are ignored. If the file is not a text file, uses a comma as the decimal mark (e.g. "20,00 -3,3", European number format), has fewer than 10 data lines, has frequencies that are not ascending or are repeated, does not include 1 [kHz], or has values beyond ±30 [dB] (absolute level data), an error message appears and the file loaded before is kept. If the data do not cover 20 ∼ 20,000 [Hz], a message says that the end values are used outside the range.

The same correction is applied to X and Y, so X − Y (red line) is the same with or without calibration. The calibration changes only the levels of the spectra and envelopes of X and Y.

■ **Format of the calibration file :** one line per frequency, [frequency (Hz) dB value] separated by spaces or a tab (the format commonly used by REW and others).

```
* Mic Model : xxxxxx xxx-xx
* Freq(Hz) dB
10.000     -1.717
10.481     -1.717
10.985     -1.717
...
23594.406  -5.440
24000.000  -5.440
```

The dB value is how much louder (+) or quieter (−) the microphone records at that frequency than a flat one. The program adds the value with its sign reversed to the spectra and envelopes. For example, if the value at 24,000 [Hz] in the file above is −5.440, +5.440 [dB] is added. Values between the frequencies of the file are interpolated on a logarithmic frequency axis. Some manufacturers supply values whose sign is already reversed, so check the description of the file.

5.6 7. ZONE / SPOT DETECTION

Judges the Zones and Spots of the result graph (X − Y) automatically (Figure 5). The rules are in section 7.

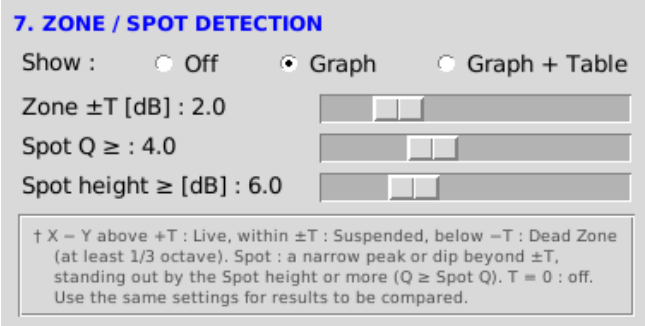


Figure 5: *7. ZONE / SPOT DETECTION*

■ **Show :** how the result is shown.

- [Off] : no judgement. Only the Zone guide (±T band) is drawn.
- [Graph] (default) : the Zone bars and Spot marks are drawn in the result graph.
- [Graph + Table] : a window with a table of the result opens as well.

■ **Zone ±T [dB] :** the value that divides the Zones (0 ∼ 10 [dB], steps of 0.5, default 2 [dB]). Green dashed lines at ±T and a light green band between them (Zone guide) are drawn in the result graph. Where the curve is above +T it is Live, within ±T Suspended, and below −T Dead Zone. At 0, Off is shown and there is neither band nor judgement. How to choose T is described in section 7.

■ **Spot Q ≥ :** how sharp a peak / dip must be to be judged a Spot (1.0 ∼ 10.0, steps of 0.1, default 4). Only narrow peaks / dips with Q at least this value are shown as Spots. Q 4 means a width of at most 1/4 of the peak frequency (section 7).

■ **Spot height ≥ [dB] :** how far a peak / dip must rise above (fall below) the curve around it (its Height) (1 ∼ 20 [dB], steps of 0.5, default 6 [dB]). Only those at least this high are shown as Spots, so if there are too many Spots, raise this value (section 7).

■ **Range of the judgement :** only the part within the Min ∼ Max range of the X-axis (section 5.4), from 20 [Hz], and outside the gray area (below f0 of a sound with harmonics) is judged.

■ **Display in the graph :** the frequency range of each Zone is shown as a thick bar just above the X-axis. The bar begins and ends at the first and last frequency of the Zone, and its color shows the kind of Zone : ▬ Live (orange), ▬ Suspended (green), ▬ Dead (black). A Live Spot is marked with an orange ▲ just above the curve and a Dead Spot with a black ▼ just below it, together with the frequency (e.g. 2.05k). The legend shows the range of each Zone (e.g. Live Zone (> +2.0 dB)) and the criteria of the Spots (e.g. Live Spot (Q ≥ 4, height ≥ 6 dB)).

■ **Result window (Graph + Table) :** from the top, the calculation information, the table of the result, and the criteria (Zone, Spot) and the meaning of each column (Figure 11(b)). In the table, the Spots of each Zone are listed (indented) in the rows below it; the meaning of the columns is given in Table 6.

Column	**Zone (Z1, Z2, ...)**	**Spot (S1, S2, ...)**
No	Z1, Z2, … : in frequency order	S1, S2, … : indented under the Zone they lie in
Type	Live / Suspended / Dead Zone	Live Spot (narrow peak) / Dead Spot (narrow dip)
From, To [Hz]	Where the Zone begins and ends	The two sides of the peak (dip), at half its Height
Peak [Hz]	–	Frequency of the top of the peak (bottom of the dip)
Level [dB]	**Median** (not the mean) of X − Y in the Zone; the usual level of the Zone	X − Y at the peak (dip)
Height [dB]	–	How far the peak rises above (the dip falls below) the curve around it; not the height above 0 [dB]
Width [oct]	Width of the Zone (octaves)	Width from From to To (octaves)
Q	–	Peak / (To − From). The higher, the narrower and sharper

Table 6: *Columns of the result*

Level and Height : Level is measured from 0 [dB], i.e. how much louder or quieter X is than Y. Height is how far the peak (dip) rises above (falls below) its surroundings and describes the shape of the curve. From the peak, the curve is followed to each side up to the first point higher than the peak, and the lowest point on each side is found; the Height is measured from the higher of these two to the peak (for a dip, the same with up and down reversed). For example, S2 in Figure 6 has Level −12.2 [dB] and Height 17.4 [dB] : at the bottom of a dip 17.4 [dB] below its surroundings (about +5 [dB]), X is 12.2 [dB] quieter than Y. The median is used for the Level of a Zone so that it is not pulled by large peaks or dips inside the Zone.

■ **Save (Save TXT / Save PDF) :** saves the contents of the result window. Both files have the same contents in the same order (Figures 6, 7).

1. Calculation information : program version, number of the graph window (#3) and pair, file names, method

and options, FFT Size, T, Q and Spot height, range of the judgement (and the f0 it starts from, if any), numbers of Zones and Spots
2. Table of the result (Table 6)
3. Criteria : summary of the rules for Zones and Spots (section 7)
4. Columns : meaning of each column (summary of Table 6)

[Save PDF] saves an A4 page with the result graph between the calculation information and the table. A long table continues on the next page, and the criteria and the meaning of the columns follow the end of the table. [Save TXT] saves a UTF-8 text file with the columns separated by tabs, so it opens directly in Excel.

The default file name has the form `SCA_X2-Y2_LPC_yyyymmdd-hhmmss`. In Refresh mode (section 5.7), pressing the same button again also replaces the result window with the new result.

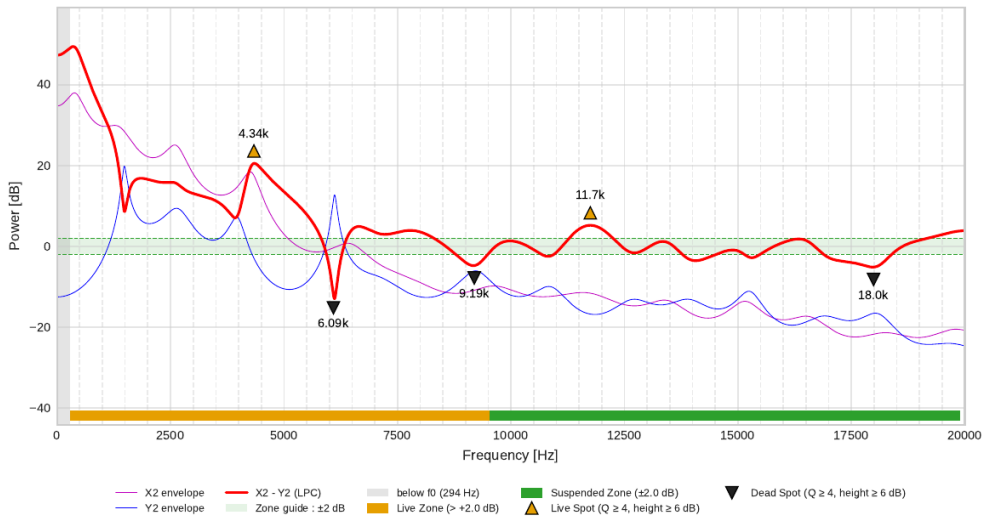
Sound Comparison Analyzer ver.4.1.0 - #3 X2 - Y2

Files : X2 = E_293.3_D3_3_0.wav, Y2 = S_293.3_D3_3_0.wav

Method : LPC, order 32, FFT 16384

T = 2 dB, Spot Q = 4, Spot height = 6 dB, judged 294 ~ 20000 Hz (from f0 294 Hz)

Zones : 2, Spots : 5 (Live 2, Dead 3)



No	Type	From [Hz]	To [Hz]	Peak [Hz]	Level [dB]	Height [dB]	Width [oct]	Q
Z1	**Live Zone**	294	9530	–	+16.7	–	5.02	–
S1	Live Spot	4125	4976	4338	+20.5	13.5	0.27	5.1
S2	Dead Spot	5969	6249	6091	-12.2	17.4	0.07	21.8
S3	Dead Spot	8557	9672	9192	-4.8	9.7	0.18	8.2
Z2	**Suspended Zone**	9530	19904	–	-0.4	–	1.06	–
S4	Live Spot	11128	12433	11750	+5.2	10.4	0.16	9.0
S5	Dead Spot	16754	18537	17990	-5.2	9.0	0.15	10.1

Criteria :

Zone	Live above +T, Suspended within ±T, Dead below −T; at least 1/3 octave wide (narrower parts are merged into the neighboring Zone).
Spot	a narrow peak (Live) or dip (Dead) : Height ≥ Spot height, Q ≥ Spot Q, above +T (Live) / below −T (Dead), and T or more beyond the Level of its Zone.

Columns :

No	Z1, Z2 ... : Zones in frequency order. S1, S2 ... : Spots, listed (indented) under the Zone they lie in.
Type	Live / Suspended / Dead Zone; Live Spot (narrow peak) / Dead Spot (narrow dip).
From, To [Hz]	Zone : where it begins and ends. Spot : the two sides of the peak (dip), at half its Height.
Peak [Hz]	Spot : frequency of the top of the peak (bottom of the dip).
Level [dB]	Zone : median of X − Y over the Zone (its usual level, not the mean). Spot : X − Y at the peak (dip).
Height [dB]	Spot : how far the peak rises above (the dip falls below) the curve around it (prominence), not the level above 0 dB.
Width [oct]	From .. To in octaves (1 = one octave, 0.5 = half an octave).
Q	Spot : Peak / (To − From). The higher, the narrower and sharper the Spot.

1 / 1

Figure 6: *Saved PDF (X2 − Y2, LPC). From the top : calculation information, result graph, table of the result, criteria (Criteria), meaning of the columns (Columns)*

```
Sound Comparison Analyzer ver.4.1.0  -  #3  X2 - Y2
Files : X2 = E_293.3_D3_3_0.wav,  Y2 = S_293.3_D3_3_0.wav
Method : LPC, order 32,  FFT 16384
T = 2 dB,  Spot Q = 4,  Spot height = 6 dB,  judged 294 ~ 20000 Hz  (from f0 294 Hz)
Zones : 2,  Spots : 5 (Live 2, Dead 3)

No    Type            From [Hz]  To [Hz]  Peak [Hz]  Level [dB]  Height [dB]  Width [oct]     Q
Z1    Live Zone             294     9530          –       +16.7            –         5.02     –
  S1  Live Spot            4125     4976       4338       +20.5         13.5         0.27   5.1
  S2  Dead Spot            5969     6249       6091       -12.2         17.4         0.07  21.8
  S3  Dead Spot            8557     9672       9192        -4.8          9.7         0.18   8.2
Z2    Suspended Zone       9530    19904          –        -0.4            –         1.06     –
  S4  Live Spot           11128    12433      11750        +5.2         10.4         0.16   9.0
  S5  Dead Spot           16754    18537      17990        -5.2          9.0         0.15  10.1

Criteria :
  Zone : Live above +T, Suspended within ±T, Dead below −T;  at least 1/3 octave wide (narrower parts are
    merged into the neighboring Zone).
  Spot : a narrow peak (Live) or dip (Dead) :  Height ≥ Spot height,  Q ≥ Spot Q,  above +T (Live) / below
    −T (Dead),  and T or more beyond the Level of its Zone.

Columns :
  No : Z1, Z2 ... : Zones in frequency order.  S1, S2 ... : Spots, listed (indented) under the Zone they lie
    in.
  Type : Live / Suspended / Dead Zone;  Live Spot (narrow peak) / Dead Spot (narrow dip).
  From, To [Hz] : Zone : where it begins and ends.  Spot : the two sides of the peak (dip), at half its
    Height.
  Peak [Hz] : Spot : frequency of the top of the peak (bottom of the dip).
  Level [dB] : Zone : median of X − Y over the Zone (its usual level, not the mean).  Spot : X − Y at the
    peak (dip).
  Height [dB] : Spot : how far the peak rises above (the dip falls below) the curve around it (prominence),
    not the level above 0 dB.
  Width [oct] : From .. To in octaves (1 = one octave, 0.5 = half an octave).
  Q : Spot : Peak / (To − From). The higher, the narrower and sharper the Spot.
```

Figure 7: *Saved TXT (X2 − Y2, LPC). In the actual file the columns of the table are separated by tabs; here the columns are aligned and long lines are wrapped for readability*

5.7 8. CALCULATION

■ **New window / Refresh :** [New window] opens a new graph window each time. [Refresh] (default) draws again in the window last drawn by the same button, so the windows do not pile up while you compare options (if that window has been closed, a new one opens). The title of a graph window shows its number, the pair, the method and its options.

■ **Calculation buttons :** [X1 − Y1] ∼ [X4 − Y4] are active (green) only for pairs whose two files are both loaded. [X all − Y all] is active (blue) when LTAS is selected and there is at least one X file and one Y file; it averages all X files and all Y files, each with equal weight, and compares them (Figure 8).

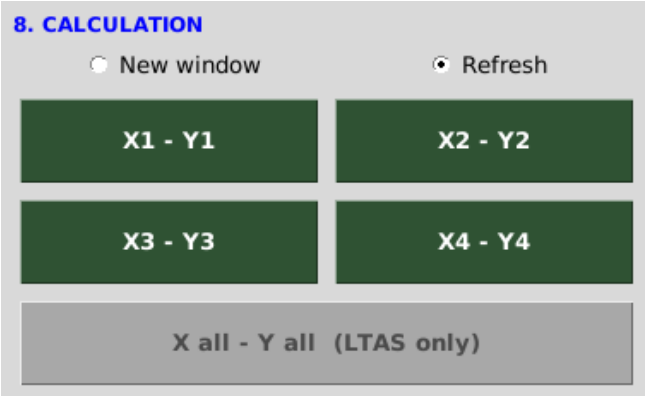


Figure 8: *Calculation buttons (8 files loaded, LPC selected)*

In the following cases a message box appears before the calculation; press [OK] to continue.

- A file shorter than the Analysis Length (the whole file is analyzed)
- The part actually analyzed (the whole file with LTAS) is shorter than the FFT Size (zero-padded and analyzed as one frame)
- Only one of X and Y has harmonics (section 7, gray area)

If files with different sampling frequencies are calculated together, or Other mic is selected without a calibration file, an error message appears and nothing is calculated. The toolbar below a graph window zooms the graph or saves it as an image file.

6 Example

An example of finding the amplification by the violin body. Load the sound of the violin with the string mounted into [SOUND SET X], and the sound of the string alone without a body into [SOUND SET Y], in the order of the G, D, A and E strings (Figure 1). All other options are at their defaults.

6.1 Comparison of one note (X1 − Y1, LPC)

Select LPC and press [X1 − Y1]; the result of Figure 9 appears.

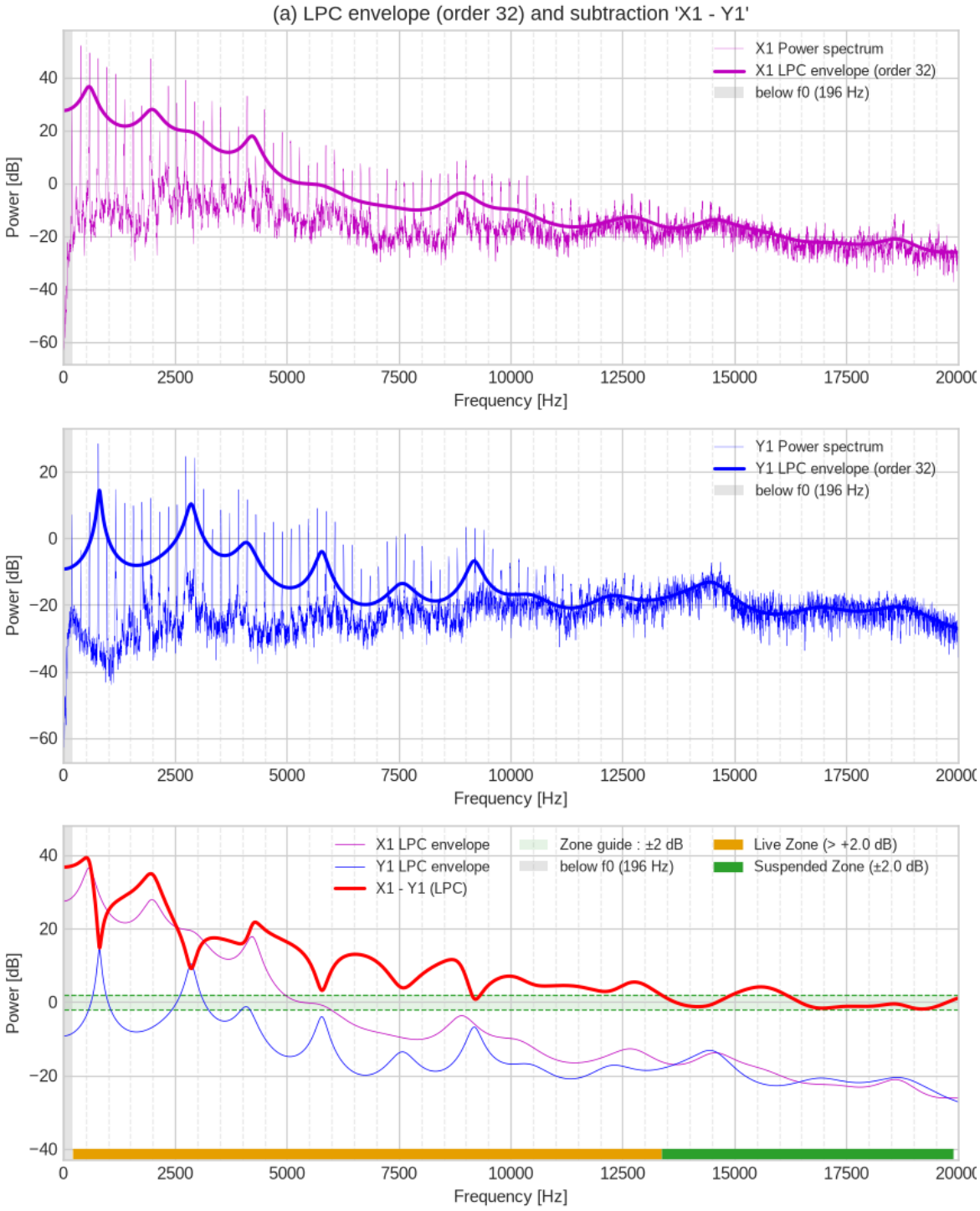


Figure 9: *Result (G string, LPC)*

■ Explanation of the result :

- The purple graph at the top is the analysis of X1 : the thin line is the power spectrum and the thick line is the envelope (LPC envelope).
- The blue graph in the middle is the analysis of Y1 (lines as above).
- In the bottom graph, the purple and blue lines are the two envelopes above, and the red line is their difference (X − Y), i.e. the result graph.

- The gray box in the legend (below f0 (196 Hz)) is the range below the fundamental of the G string sound (f0, 196 [Hz]); it is not part of this note and is not judged (section 7, gray area).
- The green band is the Zone guide (default ±2 [dB]) for the Zone judgement (sections 5.6 and 7).
- The bars just above the X-axis are the frequency ranges of the Zones found automatically. Orange is Live Zone and green is Suspended Zone; Spots, if any, are marked with an orange ▲ (Live Spot) and a black ▼ (Dead Spot) (sections 5.6 and 7).

A value of +10 [dB] in the result graph means that X1 is 10 [dB] louder than Y1 at that frequency, i.e. the sound is amplified by 10 [dB] through the body. 0 means no change, and a negative value means attenuation.

The result graph (red) in Figure 9 shows large amplification of about 35 [dB] at about 2 [kHz] and about 22 [dB] at 4.3 [kHz], and about 38 [dB] below 500 [Hz]. Around 0.8 [kHz] and 2.9 [kHz] the amplification is smaller than around them. The amplification decreases toward the treble and is almost 0 [dB] above about 13 [kHz]. The automatic judgement gives a Live Zone up to about 13.4 [kHz] and a Suspended Zone above it, and with the defaults (Spot height 6 [dB]) no Spots.

6.2 Average of several notes (X all − Y all, LTAS)

Select LTAS + Smoothing and press [X all − Y all]; the sounds of the 4 strings are averaged and compared (Figure 10). Averaging several notes makes the result less biased toward particular harmonics (this example has only the 4 open strings, so not all gaps between the harmonics are filled). The amplification is largest, about 52 [dB], at about 290 [Hz], and stands out around 1 [kHz], 2.2 [kHz] and 3.3 [kHz]; it decreases toward the treble. The automatic judgement gives a Live Zone over almost the whole range above about 76 [Hz], with a Live Spot at 0.97 [kHz]. LTAS has no gray area, so the judgement starts at 20 [Hz], but the part below the lowest note (G string, 196 [Hz]) is not part of any note and has no meaning.

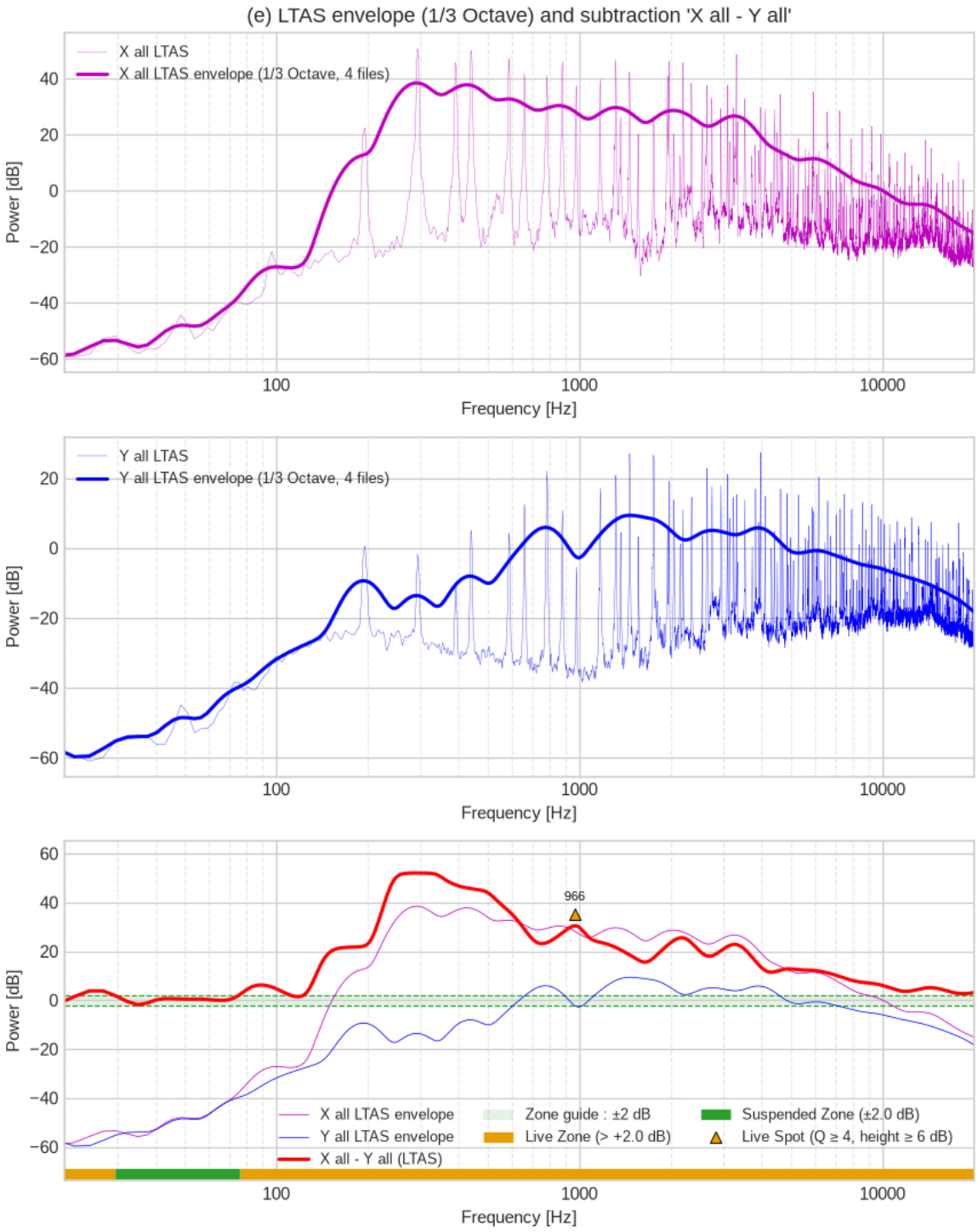


Figure 10: *Result (average of 4 strings, LTAS 1/3 Octave, Log scale)*

7 Reading the result graph

The result graph (red line, X − Y) is read by dividing it into the 5 types of Table 7. With Show set to Graph or Graph + Table, the program judges them automatically by the rules below (section 5.6).

Kind	Name	Rule	Meaning
Zone	**Live Zone**	Above +T	Amplified range
	Suspended Zone	Within the ±T band	Neither amplified nor attenuated
	Dead Zone	Below −T	Attenuated range
Spot	**Live Spot**	Narrow peak, above +T	Local amplification
	Dead Spot	Narrow dip, below −T	Local attenuation

Table 7: *5 types of the result graph (every Zone is at least 1/3 octave wide)*

■ **Judgement of the Zones :** a Zone shows, over a range of frequencies, whether X has become louder than Y (Live), quieter (Dead), or shows no difference (Suspended). Each point of the curve is classed as above +T, within ±T or below −T; then parts narrower than 1/3 octave are merged into a neighboring Zone, the narrowest first (into the kind of both neighbors if they are the same, otherwise into the wider one). So short excursions out of the band do not split a Zone; such parts are judged as Spots.

T sets how large a difference still counts as "no difference". Even the same sound recorded again under the same conditions usually gives results that differ by $1 \sim 2$ [dB] because of small differences in playing strength, microphone position and so on, so the default is 2 [dB]. If you record the same sound twice and compare the two as X and Y, the width by which the result graph moves up and down around 0 is the error of the measurement, and T can be set to that width. Results to be compared must be calculated with the same T.

■ **Judgement of the Spots :** a Spot is a narrow place where X has become louder (Live) or quieter (Dead) than Y. Of the peaks (Live Spot candidates) and dips (Dead Spot candidates) of the result graph, only those that meet all four conditions are judged Spots.

1. Height : rises above (falls below) its surroundings by at least [Spot height] (default 6 [dB])
2. Sharpness : Q at least [Spot Q] (Q = peak frequency / width, the width measured at half the Height)
3. Real change : a Live Spot's peak is above +T, a Dead Spot's dip is below −T
4. Difference from the Zone : at least T above (below) the usual level of its Zone (the median of X − Y)

Condition 3 means that a Spot, like a Zone, must be a place where X has really become louder or quieter. For example, a peak in a Dead Zone that does not reach +T is only "less attenuated" and is not a Live Spot, and a dip in a Live Zone that stays above −T is not a Dead Spot either. Without condition 4, a plain valley between the end of a Zone and a high neighboring peak would be judged a Dead Spot (its depth and Q are set by the heights and distance of its neighbors, so Height or Q alone cannot remove it). When a Spot takes more than half of its Zone, e.g. when a large peak has been judged a Zone by itself, the median of ±1 octave around the Spot is used instead of the Zone.

The width is measured at half the Height because a small peak of $2 \sim 3$ [dB] may have no point at −3 [dB]. For a Height of 6 [dB] or more, this width equals the −3 [dB] bandwidth, and Q matches the usual definition (peak frequency / −3 [dB] bandwidth).

Since width = peak frequency / Q, Spot Q sets the largest width judged as a Spot, as a fraction of the peak frequency (Table 8).

Q	Width / peak frequency	Width [octaves]	e.g. width of a 1 [kHz] peak [Hz]
1.4	1/1.4 (71%)	1.00	710
2	1/2 (50%)	0.71	500
3	1/3 (33%)	0.48 (≈ 1/2)	333
4 (default)	**1/4 (25%)**	**0.36 (≈ 1/3)**	**250**
5	1/5 (20%)	0.29	200
8.7	1/8.7 (11.5%)	0.17 (≈ 1/6)	115
10	1/10 (10%)	0.14	100

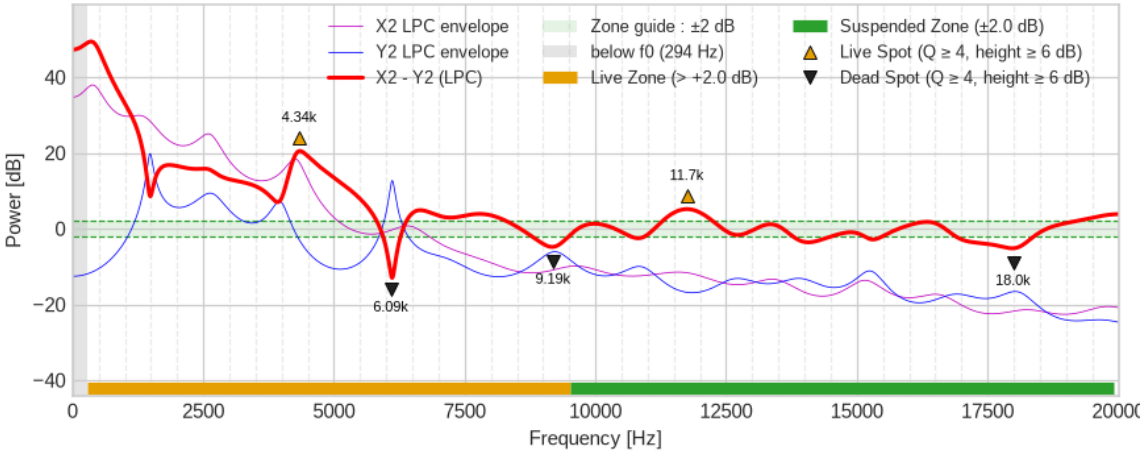
Table 8: *Spot Q and the width of a Spot (only those up to this width are judged Spots)*

Why Q 4 and Spot height 6 [dB] : Q 4 is a width of 1/4 of the peak frequency, about 1/3 octave. 1/3 octave is the minimum width judged as a Zone, so peaks and dips narrower than this are judged as Spots. A Spot height of 6 [dB] means twice the amplitude of the surroundings, so only clearly prominent ones remain. The shallow ripples of a high-order LPC and the like are removed by this even when their Q is high, because they are low. Lowering Q or Spot height lets wider or shallower ones be judged as Spots, and the number of Spots increases. Results to be compared must be calculated with the same Q and Spot height, as with T.

The judgement of Zones and Spots is meaningful only when X and Y were recorded at the same recording level (gain, distance).

■ **Example of the judgement :** Figure 11 is the result of X2 − Y2 (D string, LPC) of section 6, calculated with Show [Graph + Table]. There is a Live Zone up to about 9.5 [kHz] (median +16.7 [dB]) and a Suspended Zone above it. In the Live Zone there are a Live Spot at 4.34 [kHz] (+20.5 [dB], Height 13.5 [dB], Q 5.1), a distinct Dead Spot at 6.09 [kHz] (−12.2 [dB], Height 17.4 [dB], Q 21.8) and a Dead Spot at 9.19 [kHz]; in the Suspended Zone there are a Live Spot at 11.7 [kHz] and a Dead Spot at 18.0 [kHz]. The small dips around 10.9 [kHz] and 14.1 [kHz] go below −T, but their Heights of 3.8 [dB] and 5.4 [dB] do not reach the Spot height (6 [dB]). The dip around 1.5 [kHz] falls deeply, but its value of about +9 [dB] is still amplified, so it is not a Dead Spot (condition 3).

X1 − Y1 of section 6.1 (Figure 9) has a Live Zone up to about 13.4 [kHz] and a Suspended Zone above it. The short rise above the band around 15.6 [kHz] has a Height of 5.4 [dB] and is not a Spot with the defaults; with a Spot height of 5 [dB] it is judged a Live Spot.



(a) *Result graph (Zone bars and Spot marks)*

Sound Comparison Analyzer ver.4.1.0 - #3 X2 - Y2
Files : X2 = E_293.3_D3_3_0.wav, Y2 = S_293.3_D3_3_0.wav
Method : LPC, order 32, FFT 16384
T = 2 dB, Spot Q = 4, Spot height = 6 dB, judged 294 ~ 20000 Hz (from f0 294 Hz)
Zones : 2, Spots : 5 (Live 2, Dead 3)

No	Type	From [Hz]	To [Hz]	Peak [Hz]	Level [dB]	Height [dB]	Width [oct]	Q
Z1	Live Zone	294	9530	-	+16.7	-	5.02	-
S1	Live Spot	4125	4976	4338	+20.5	13.5	0.27	5.1
S2	Dead Spot	5969	6249	6091	-12.2	17.4	0.07	21.8
S3	Dead Spot	8557	9672	9192	-4.8	9.7	0.18	8.2
Z2	Suspended Zone	9530	19904	-	-0.4	-	1.06	-
S4	Live Spot	11128	12433	11750	+5.2	10.4	0.16	9.0
S5	Dead Spot	16754	18537	17990	-5.2	9.0	0.15	10.1

Zone Live above +T, Suspended within ±T, Dead below −T; at least 1/3 octave wide (narrower parts are merged into the neighboring Zone).
Spot a narrow peak (Live) or dip (Dead) : Height ≥ Spot height, Q ≥ Spot Q, above +T (Live) / below −T (Dead), and T or more beyond the Level of its Zone.

No Z1, Z2 ... : Zones in frequency order. S1, S2 ... : Spots, listed (indented) under the Zone they lie in.
Type Live / Suspended / Dead Zone; Live Spot (narrow peak) / Dead Spot (narrow dip).
From, To [Hz] Zone : where it begins and ends. Spot : the two sides of the peak (dip), at half its Height.
Peak [Hz] Spot : frequency of the top of the peak (bottom of the dip).
Level [dB] Zone : median of X − Y over the Zone (its usual level, not the mean). Spot : X − Y at the peak (dip).
Height [dB] Spot : how far the peak rises above (the dip falls below) the curve around it (prominence), not the level above 0 dB.
Width [oct] From .. To in octaves (1 = one octave, 0.5 = half an octave).
Q Spot : Peak / (To − From). The higher, the narrower and sharper the Spot.

Save TXT | Save PDF | Close

(b) *Result window*

Figure 11: *Automatic judgement of Zones and Spots (X2 − Y2, LPC, T = 2 [dB], Q = 4, Spot height = 6 [dB])*

Even for the same sound, the result graph changes with the kind of envelope and its options (order, Smoothing, etc.) and the FFT settings (FFT Size, Analysis Length). Choose the options that suit the purpose (section 5.2, How to choose a method), and always calculate results to be compared with the same options.

■ **Gray area (below f0) :** a sound to be analyzed may have harmonics (instrument tones, etc.) or not (tap tones, impulses, sweeps, noise, etc.). The program first checks whether the sound has harmonics, and only for a sound with harmonics paints the part below the fundamental (f0) gray and leaves it out of the judgement. A sound without harmonics has no f0, so its whole range is compared as it is, with no gray area.

For a sound with harmonics, the part below f0 has no meaning for the following reason. For example, a violin playing A4 consists of 440 [Hz] (f0) and its harmonics at integer multiples (880 [Hz], 1320 [Hz], …), so the

lowest component of that note is 440 [Hz]. Below 440 [Hz] there may be bow noise, attack transients, room sound and so on, but they are not part of the note and cannot be used to compare two notes. In addition, LPC, True Envelope and MGCE are envelopes that connect the harmonic peaks, so below f0 the curve is not a measurement but stays near the height of the f0 peak (in LPC it is the tail of the resonance model fitted to the harmonic peaks; in True Envelope and MGCE the spectrum is treated as symmetric about 0 [Hz] in the cepstrum calculation, so the curve bridges the peaks at −f0 and +f0). The Smoothing curve below f0 is a measurement, but it is not part of the note, so it is grayed by the same rule.

The upper and middle graphs use the f0 of X and Y respectively, and the result graph uses the higher of the two; if only one has harmonics, its f0 is used. LTAS (the average of many notes) has no single f0 and is not grayed, but for the same reason the part below the f0 of the lowest note has no meaning.

Harmonics are checked as follows. A candidate f0 is found from the periodicity of the sound (30 ∼ 4500 [Hz]), and the program looks for harmonic peaks at f0, 2f0, … 8f0. If at least 5 of these 8 stand 10 [dB] or more above their surroundings and lie within 60 [dB] of the largest peak, the sound is regarded as having harmonics. If the FFT Size is too small to separate the harmonics (e.g. FFT Size 2048 for a low note), the sound is treated as having no harmonics and no gray area is drawn.

If only one of X and Y has harmonics (e.g. a played note and a tap tone), a message box appears before the calculation. In a sound with harmonics the energy is concentrated in the harmonics, while in a sound without them it is spread out, so the levels of the two sounds change differently with the FFT Size. The height of the whole X − Y curve therefore depends on the FFT Size; only the shape of the curve (where it rises and falls) can be compared, and no Zone judgement from 0 [dB] is possible. If this message appears although both are instrument tones, the f0 of one of them was not found; try a larger FFT Size.

8 Main changes

- Program renamed : Sound Comparing Analyzer → Sound Comparison Analyzer
- Automatic judgement of Zones and Spots : new item 7. ZONE / SPOT DETECTION (Zone ±T, Spot Q, Spot height, Show : Off / Graph / Graph + Table), Zone bars and Spot marks in the result graph
- Result window : table of Zones and Spots, saved as TXT / PDF
- Screen layout : 4. FFT SETTINGS moved to the left column, calculation now item 8, descriptions on the screen shortened to the essentials
- Defaults : graph window Refresh, Show Graph
- Fixed an error when drawing again with Refresh into a closed graph window
- Envelope methods : True Envelope, Smoothing, MGCE and LTAS added to LPC (5 methods)
- [X all − Y all] : comparison of the average of all loaded files (LTAS)
- Frequency analysis : the whole Analysis Length is averaged in frames (before, only the FFT Size was used); DC component removed
- Analysis Length up to 5.0 s; status line and message boxes added
- Microphone calibration (EMM-6 built in; import and check of calibration files of other microphones)
- 8/16/24/32 bit and floating point wav files and stereo files supported (converted to mono)
- Gray area below the fundamental (f0) (only for sounds with confirmed harmonics); graph window New window / Refresh
- Zone guide (±T [dB]) in the result graph
- New screen layout (2 columns), button colors showing their state, resizable and scrollable window, sliders

adjustable with the mouse wheel

9 Development and environment

This program was developed with 'Python 3.11 + Tkinter' (NumPy, SciPy, Matplotlib) and tested on 'Windows 11'.