Zwicker Loudness Measurement with R&S® UPV Audio Analyzer
Application Note

Products:

- R&S®UPV
- R&S®UPV-K1

This application note provides an application program for the audio analyzer R&S®UPV which calculates and displays the loudness of non-stationary sounds according to ISO 532-1 (Zwicker method). The theoretic background of the evaluation method is briefly explained.

Note:

Please find the most up-to-date document on our homepage http://www.rohde-schwarz.com/appnote/1GA67.

This document is complemented by software. The software may be updated even if the version of the document remains unchanged

Application Note
T. Lechner, A. Pfeiffer
9.2017 – 0e

ROHDE&SCHWARZ

Table of Contents

1 Introduction

At first hand, measurement instruments capture physical parameters like voltage versus time. Such signals can be described by properties like time function, spectrum or RMS (root mean square) voltage. Other common analysis methods for audio signals are THD (total harmonic distortion) or SINAD (signal to noise and distortion ratio). Such properties are very useful for characterizing technical devices like amplifiers and loudspeakers.

By using transducers, other physical entities like for example force, acceleration or sound pressure can be transformed to a proportional voltage and therefore be measured.

There are, however, other measurement tasks, where the focus is on human perception. For example, in mobile communication the perceived quality of transmitted speech is of interest.

Another basic auditory sensation is loudness. The loudness of unwanted noise has an influence on its annoyance, and the loudness of a ringer tone of a mobile phone gives a hint how well an incoming call will be recognized in noisy environment.

Physical magnitudes and perceptive magnitudes	
Physical magnitude	**Perceptive magnitude**
Sound intensity	Loudness
Frequency	Pitch
AM modulation depth	Fluctuation strength
AM modulation frequency	Roughness
Spectral energy distribution	Sharpness
Duration	Subjective duration
	Speech quality
	Sound quality

Table 1-1: Physical magnitudes and perceptive magnitudes

A first approach for determining perceptive magnitudes like speech quality or loudness is subjective judgement by a group of listeners. By employing a large group and averaging the judgements, a reasonable repeatability of the scores can be achieved. However, there will always be a variation if the same sound is judged by a different group. Furthermore, listening tests are costly because human labor is involved.

Repeatability and effectivity are motivation for determining analysis methods which can predict human judgement for a certain perception. Such analysis methods have to take the properties of human perception into account.

Loudness calculation according to Eberhard Zwicker, which is standardized in ISO 532-1 (ISO, 2017-06) and DIN 45631-A1 (DIN, 2010-03), is commonly acknowledged as instrumental method for determining perceived loudness of stationary and time-variant sound. This application note provides an application program for the R&S® UPV audio analyzer, herein below called “UPV”, and gives a short overview over the theory behind the method.

2 Theoretical background

2.1 Psychoacoustics

The science of psychoacoustics employs engineering methods and terminology for describing the properties of the human hearing. This chapter of the application note is not intended to replace a textbook. It only gives a short idea of the theory behind loudness calculation. For more detailed information please see (E. Zwicker, 1990).

The original motivation for psychoacoustic research had been judgement of audibility of auditory impairments introduced by transmission and storage, for example the idle noise of magnetic tape recording or “wow and flutter” (frequency modulation introduced by imperfections of mechanic storage). With increasing psychoacoustic knowledge and advancing technology, new applications became feasible like MP3 codecs for data reduction on music recordings by irrelevance reduction, i.e. removing information about non-perceivable features of the sound.

Listening experiments are a prominent method of psychoacoustic research. Observers listen to sound samples or pairs of sound samples with a certain task. A simple example would be to first play a complex sound with a defined level and then a sine tone, with the task for the observer to adjust the level of the sine tone such that it is perceived to have the same loudness as the complex sound. In another experiment the observer could be presented the same sound twice with the task to adjust the second instance in level such that it is perceived twice as loud as the first one.

2.1.1 Frequency analysis in the human ear

The human ear can be divided in outer ear (acoustic part up to the ear drum), middle ear (mechanic part between ear drum and oval window) and inner ear (hydromechanic part behind the oval window). Changes in air pressure excite vibrations in the ear drum which are transferred by the middle ear ossicles to the oval window. The middle ear ossicles constitute a kind of mechanic transformer which translates large excursion with low force at the ear drum to small excursion with larger force at the oval window.

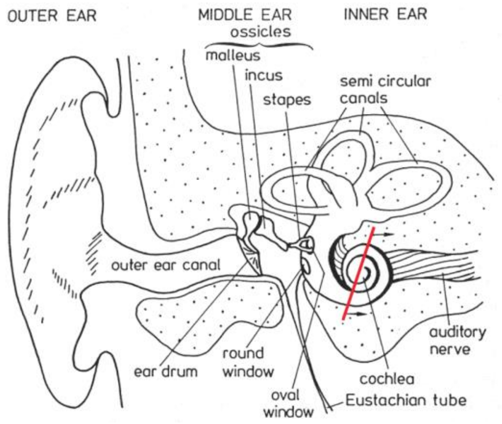

Fig. 2-1: Physiology of the human ear, from (E. Zwicker, 1990)

Fig. 2-2 shows a cross-section through the inner ear at the plane marked in red in Fig. 2-1. The hydromechanics of the cochlea consists mainly of two fluid-filled longitudinal cavities, the scala vestibuli and the scala tympani, which are separated by the basilar membrane. The basilar membrane is narrow near the entrance of the cochlea (oval window and round window) and wide near the end of the cochlea. Scala vestibuli and scala tympani are connected at the apex of the cochlea by an opening called helicotrema.

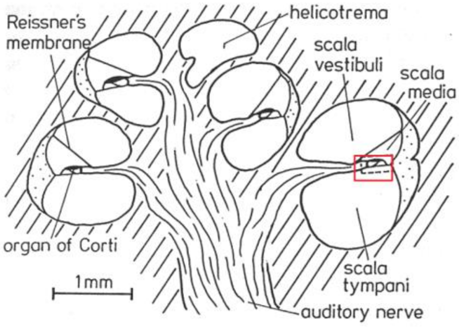

Fig. 2-2: Cross-section through the inner ear at the plane marked in red in Fig. 2-1, from (E. Zwicker, 1990)

The vibration of the stapes moves the membrane in the oval window which in turn moves the fluid in the scala vestibuli in and out. As fluid is almost incompressible, the round window moves out by the same volume as the oval window moves in, and vice versa. For very low frequencies, the volume balancing between scala vestibuli and scala tympani takes place at the helicotrema. With rising frequency the inertial mass of the fluid poses a rising impedance against the movement, and the volume balancing between scala vestibule and scala tympani takes place by movement of the basilar membrane. The consequence is a maximum of vibration of the basilar membrane which is close to the helicotrema for low frequencies and close to the oval window for high frequencies.

Fig. 2-3 shows a cross-section through the basilar membrane near the base of the cochlea.

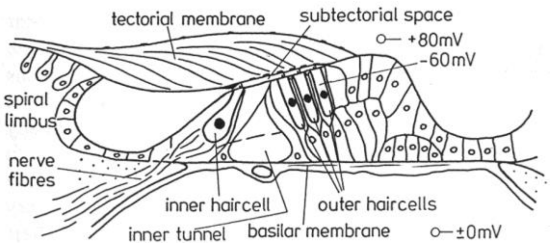

Fig. 2-3: Cross-section through the basilar membrane in the plane marked in red in Fig. 2-2, from (E. Zwicker, 1990)

The inner hair cells are sensory cells, which detect the vibration amplitude at their respective place on the basilar membrane. The outer hair cells have motoric functionality, which “amplifies” the vibration of the basilar membrane.

The “frequency response” of an inner hair cell at a certain place on the basilar membrane (tuning curve) can be described as an asymmetric band-pass with a steeper slope towards higher frequencies. The non-linear amplification effected by the outer hair cells increases the resonance at the center frequency of the band-pass curve and thus improves the frequency selectivity. Non-linear means that the amplification decreases with rising amplitude. The signal processing in the inner ear can be regarded as a set of such band-pass filters with a strong overlap. The bandwidth of these band-pass filters, called “critical bandwidth”, is approximately 100 Hz for center frequencies up to 500 Hz and approximately 20% of the center frequency for center frequencies above 500 Hz.

For loudness calculation purpose the overlap is neglected, and the frequency analysis of the ear is modeled by a filter bank of symmetric filters, arranged on a frequency scale without spectral gaps. The channels of the filter bank are called “critical bands”, and the number of adjacent critical bands up to a given frequency is called “critical band rate”. The unit for critical band rate is Bark.

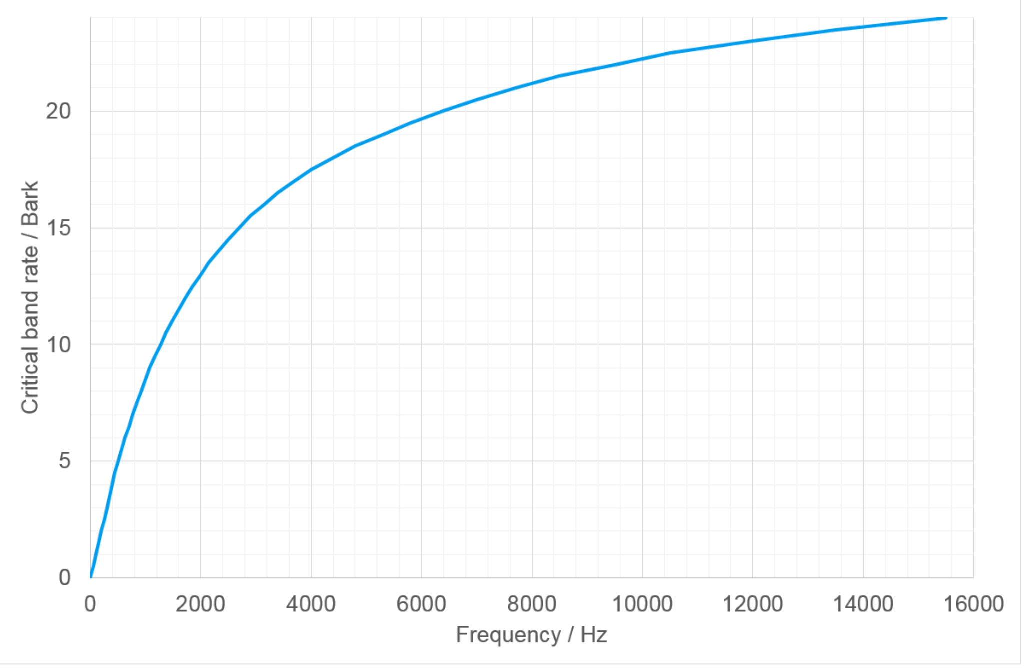

Fig. 2-4: Critical band rate as a function of frequency according to (E. Zwicker, 1990)

2.1.2 Sound pressure and sound intensity

In silence, the air in the atmosphere has a static DC pressure of about 100,000 Pa. A periodically moving membrane like a loudspeaker moves air molecules in front of the membrane (sound velocity). The movement of air molecules causes small AC changes of the air pressure which propagate into the surroundings. The sound pressure p is the RMS value of this AC component.

Sound pressure level (SPL) is the logarithmic ratio of the sound pressure referenced to the average threshold of hearing in quiet at 1 kHz, which is 2*10^{-5} Pa = 20 µPa.

$$SPL = 20 * log_{10}\left(\frac{p}{20\ \mu Pa}\right)$$

Sound intensity I is the power per unit area which passes through a certain boundary like for example the entrance of the human ear canal. Sound intensity is proportional to the square of the sound pressure, and inversely proportional to the radiation impedance of the propagating wave, which in turn depends on the shape of the wave.

In a plane wave propagating in one dimension of space, the reference sound pressure of 20 µPa corresponds to a sound intensity I_0 of 10^{-12} W/m^2. A plane wave may be regarded as the sound radiated from an infinitely large moving membrane.

2.1.3 Loudness as perceptive magnitude for sound intensity

A scale of loudness depending on sound intensity can be produced by alternately presenting two sounds of identical spectral distribution (e.g. 1 kHz tone) with the task for the observer to adjust the level of the second sound such that is perceived twice as loud as the first (reference) one. The unit for loudness is “sone”, and the loudness of a 1 kHz tone with an SPL of 40 dB is defined as 1 sone.

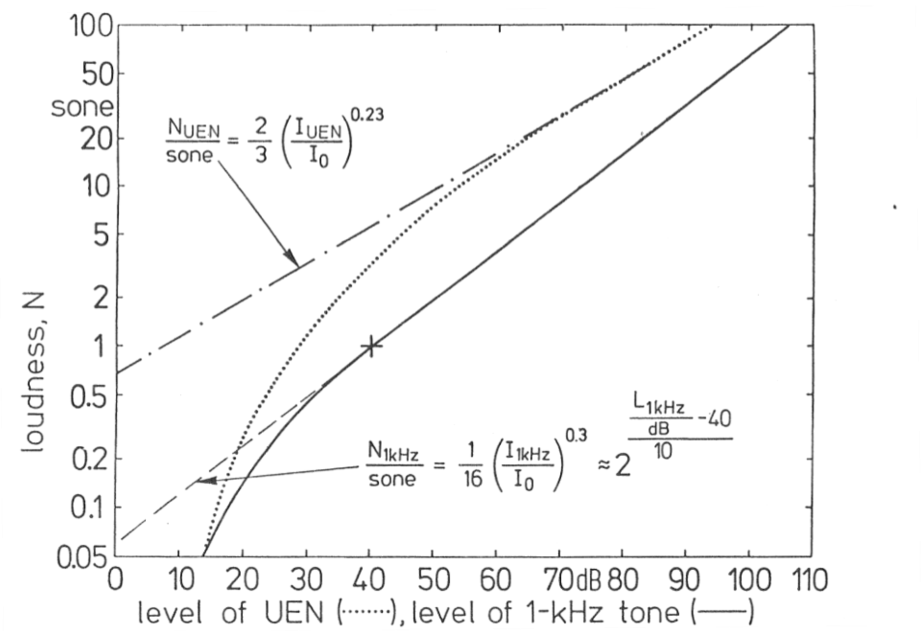

Fig. 2-5: Loudness function of a 1 kHz-tone (solid line) and of broadband “uniform-exciting” noise (dotted). From (E. Zwicker, 1990).

Fig. 2-5 shows the result for a sine tone (1-kHz tone) and for broadband noise (UEN).

For a short duration below 300 ms, the perceived loudness of a sound decreases with decreasing duration.

2.1.4 Masking

The term “masking” describes the phenomenon that a certain sound (e.g. “test tone”) which is audible with no other sound present, becomes inaudible by the presence of another sound (masker). A typical example in everyday life is that an incoming call is missed while walking in a street, because the ring tone is masked by the street noise.

The term “spectral masking” or “simultaneous masking” describes the masking effect of sounds simultaneously present with the masked sound. Masking can also occur if a loud sound precedes (“post-masking”) or even follows (“pre-masking”) the masked sound.

If the masker is not strong enough for completely rendering the masked sound inaudible, the loudness of the masked sound is reduced. This effect is called "partial masking".

2.2 Loudness Calculation for stationary sound according to ISO 532-1 Section 5 and DIN 45631

2.2.1 Overview

The following figure shows a block schematic of the loudness calculation for stationary sound.

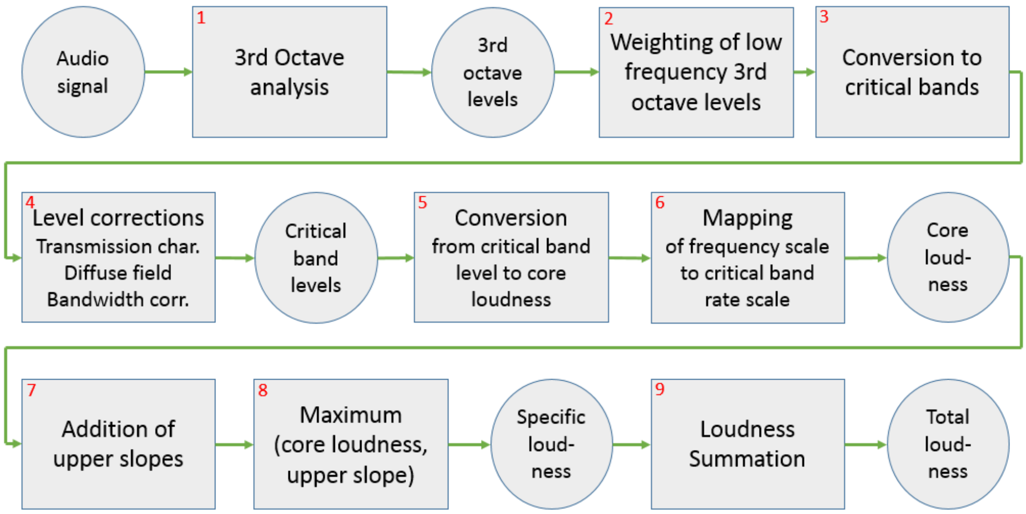

Fig. 2-6: Block schematic of the loudness calculation process for stationary sound

1. 3rd octave analysis with center frequencies from 25 Hz to 12.5 kHz.
2. Weighting of low-frequency 3rd octave levels to account for the lower sensitivity of the ear at low frequencies.
3. Combination of energy from 3rd octaves at low frequencies:
   25 Hz, 31.5 Hz, 40 Hz, 50 Hz, 63 Hz and 80 Hz
   100 Hz, 125 Hz and 160 Hz
   200 Hz and 250 Hz.
4. Corrections of the critical band levels for the transmission characteristics of the human ear, for diffuse-field sound (if applicable) and for the deviation of the frequency bands resulting from the 3rd octave analysis from the actual critical bandwidths. As critical bands are approximated by third octave bands or (at low frequencies) combinations of third octave bands, the level in each band caused by broadband sound deviates from the level in the corresponding actual critical band, which is corrected in this step.
5. Conversion of critical band levels to core loudness. The core loudness represents the signal level in the critical band on a psychoacoustic scale, albeit

without taking the slopes into account which result from the asymmetry of the frequency selectivity in human hearing. The latter is done in step 7.

6. Mapping of the frequency scale to the critical band scale yields the core loudness vs. critical band rate.
7. Addition of slopes towards higher critical bands to account for asymmetry of the frequency analysis in the ear. The steepness of the slopes depends on the core loudness in the respective critical band (see calculation example in 2.2.2).
8. Selection of the maximum out of (a) slope and (b) core loudness in each critical band yields the specific loudness vs. critical band rate.
9. Integration of the specific loudness over the critical band rate scale to yield the total loudness.

2.2.2 Calculation example for stationary sound

For manually calculating the loudness of stationary sounds there is a graphical method using templates from (DIN, 1991-03).

As an example the loudness of fan noise in a room shall be calculated.

f/Hz	20	25	31.5	40	50	63	80	100	125	160
SPL/dB	33	32	28	28	26	22	25	23	28	24

f/Hz	200	250	315	400	500	630	800	1000	1250	1600
SPL/dB	23	23	32	24	22	29	20	23	27	22

f/Hz	2000	2500	3150	4000	5000	6300	8000	10000	12500
SPL/dB	20	25	21	19	23	19	18	19	17

Table 2-1: 3rd octave levels for calculation example

Step 3: Combining energies from multiple 3rd octave intervals at low frequencies:

$$L_1 = 10 * log_{10}\left(10^{\frac{33}{10}} + 10^{\frac{32}{10}} + 2 * 10^{\frac{28}{10}} + 10^{\frac{26}{10}} + 10^{\frac{22}{10}} + 10^{\frac{25}{10}}\right) dB = 37.6\ dB$$

$$L_2 = 10 * log_{10}\left(10^{\frac{23}{10}} + 10^{\frac{28}{10}} + 10^{\frac{24}{10}}\right) dB = 30.3\ dB$$

$$L_3 = 10 * log_{10}\left(2 * 10^{\frac{23}{10}}\right) dB = 26\ dB$$

Steps 2 and 4-6 are incorporated in the design and scaling of the template. There are different templates for free sound field and diffuse sound field. The template for diffuse sound field incorporates the diffuse field correction.

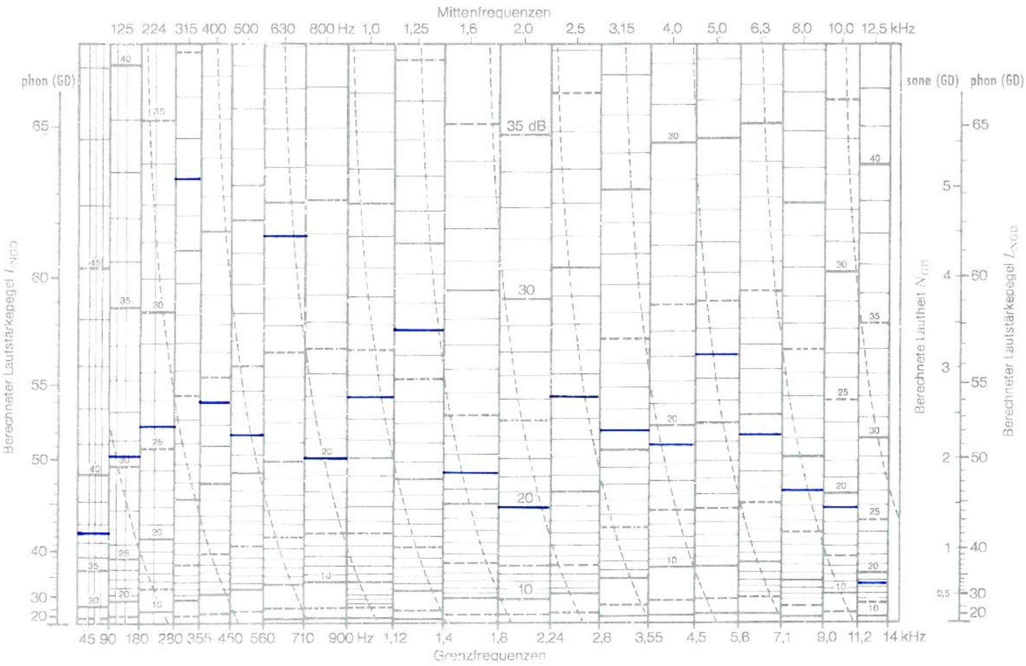

Fig. 2-7: Loudness calculation template from (DIN, 1991-03) with core loudness (blue)

The level for each frequency band is entered with a horizontal line according to the scale in the respective band. The height of each blue line represents the core loudness in this band.

Slopes are attached to the core loudness. Rising slopes are vertical. Falling slopes are entered “in parallel” to the dashed lines, i.e. with constant horizontal distance.

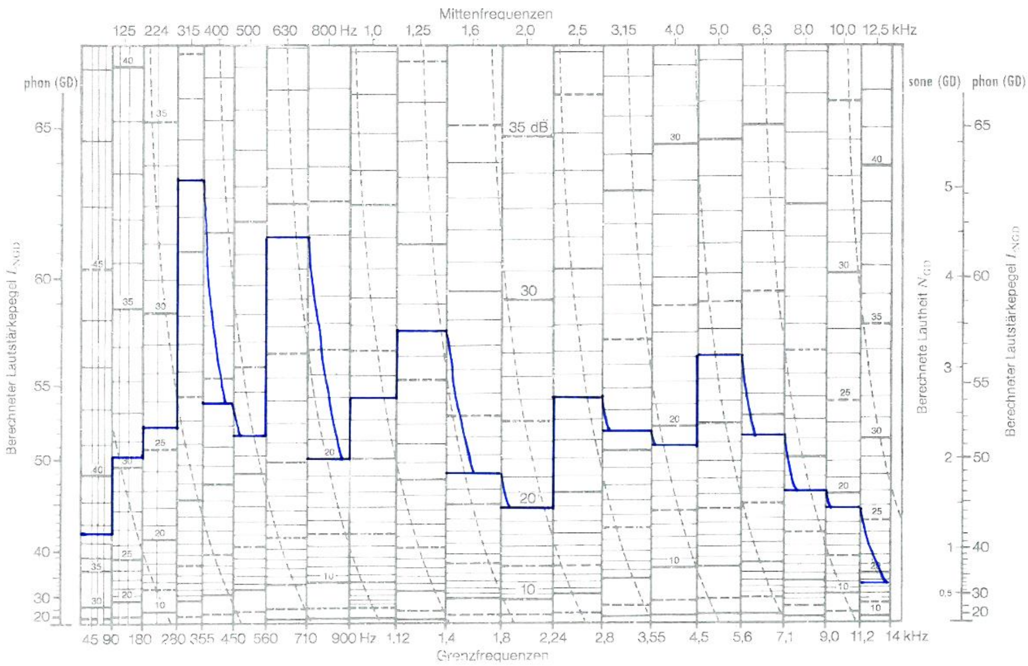

Fig. 2-8: Loudness calculation template from (DIN, 1991-03) with specific loudness (blue)

At 400 Hz, 800 Hz and 12.5 kHz the core loudness does not contribute significantly to the overall loudness because it is lower than the slope of the adjacent band. The signal in this frequency range is masked by spectral components at lower frequencies.

The size of the area below the blue curve represents the total loudness. The size of the area can be estimated by finding a straight line which averages the blue line, i.e. the area below the blue line is approximately equal to the area below the straight line.

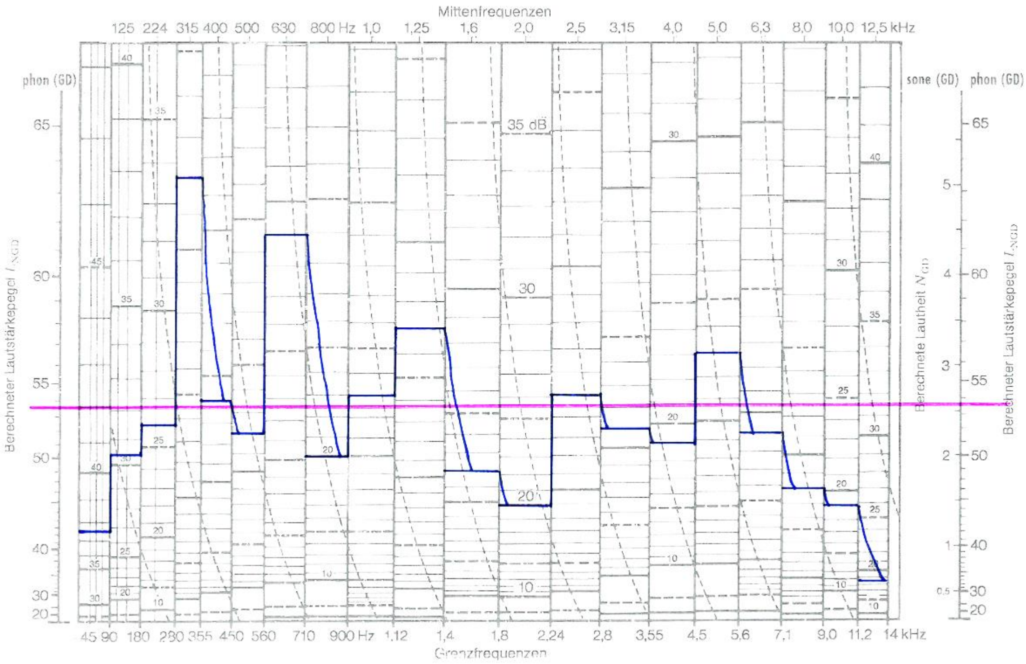

Fig. 2-9: Loudness calculation template from (DIN, 1991-03) with estimated total loudness

The total loudness can be read where the horizontal line crosses the vertical scale. In the present example the calculated loudness is about 2.5 sone, and the calculated loudness level is 53.5 phon.

2.3 Loudness calculation for non-stationary sound according to ISO 532-1 Section 6 and DIN 45631/A1

When evaluating a stationary sound, the loudness calculation method for non-stationary sound delivers the same result as the method described in section 2.2. In addition to the steps described there, the evaluation method for non-stationary sound comprises some further steps for modelling the temporal effects of loudness.

The core loudness versus time in each frequency band is subject to nonlinear decay, in which the falling slope after short peaks in the core loudness is steeper than the falling slope after longer lasting periods of high core loudness.

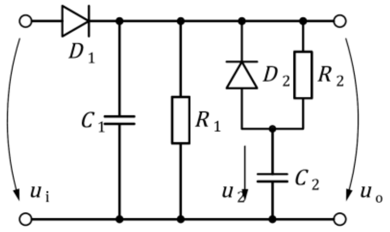

Fig. 2-10: Schematic simulating the nonlinear decay of the core loudness (ISO, 2017-06)

The total loudness is subject to a final temporal weighting consisting of two low-pass filters with different time constant, whose output signals are added with certain weighting factors.

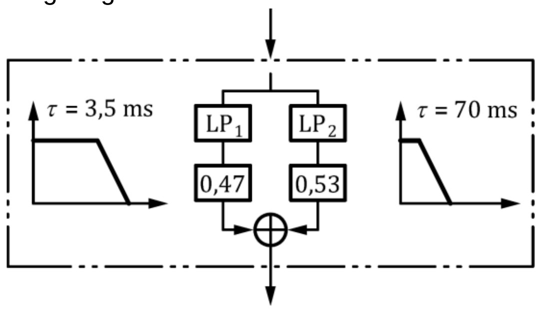

Fig. 2-11: Structure for the temporal weighting of the total loudness (ISO, 2017-06)

Besides the temporal weighting, the method for nonstationary sounds performs a “stationary” loudness calculation in regular time intervals. The RMS detection following the 3rd octave filters runs continuously. The resulting RMS values are downsampled in order to reduce calculation effort.

Due to the low-pass characteristics of the loudness perception, it is sufficient to calculate a new total loudness value every two milliseconds. Only the nonlinear decay and the final low-pass filters are calculated at a higher sample rate of 48 kHz for enhanced precision.

2.4 Percentile evaluation

Percentiles are a method of statistic evaluation. For a quantity of z different values V_i, a value V_x is determined such that $V_i \geq V_x$ for a number $z * x / 100$ values. Note that $i \in [1..z]$ is the index of a single value, whereas x denotes the percentage defining the division of all V_i in a subset below V_x and a subset above V_x. In the case of loudness evaluation, V_i are the loudness values N_i calculated versus time. V_x is called percentile loudness N_x. For example, N_5 is the loudness value which is met or exceeded in 5% of the measurement time.

The total loudness values versus time are subject to percentile evaluation in order to determine the overall perceived loudness of time-variant sound like speech or noise, e.g. from vehicles randomly passing by.

According to (ISO, 2017-06), percentile loudness N_5 shall be used to determine the overall perceived loudness. For determining the overall loudness of speech, N_7 may be more appropriate.

3 Preparations

3.1 Prerequisites

Running the application program requires option R&S® UPV-K1 (universal sequence controller).

3.2 Installation of the application program

Copy the installation file “1GA67_130.msi” to the UPV hard disk and start it. The installer creates shortcuts on the desktop and in the Programs menu. After the first start of the application program, a working folder “D:\UPV\Zwicker-Loudness” and a folder “D:\UPV\FilterData” with auxiliary data are created.

4 User interface of the application program

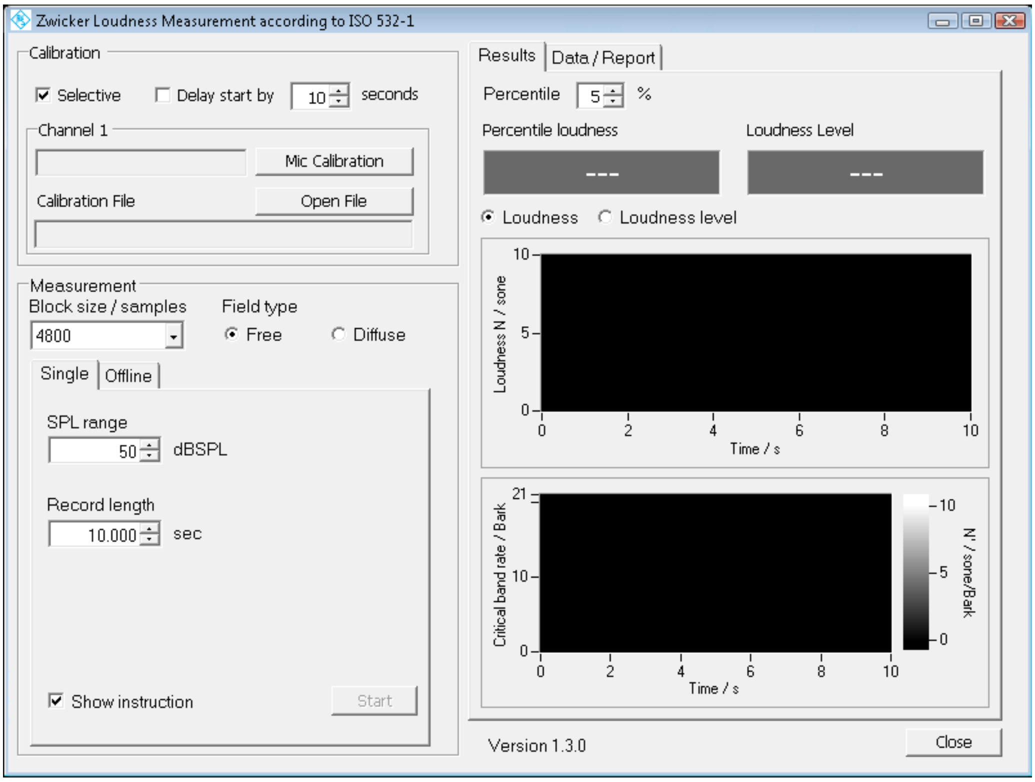

Fig. 4-1: Main window of the application program

4.1 Microphone calibration

Either a new microphone calibration can be performed or an already existing microphone calibration can be loaded. Selective calibration is recommended. It reduces the influence of broadband background noise in the calibration process. In the case of selective calibration, the frequency of the calibration tone generated by the acoustic calibrator has to be entered. If the calibration should start with a delay, mark “Delay start by” and enter the time you need in seconds in the numeric up down.
To start a new calibration press “Mic Calibration” and follow the instructions. After successful microphone calibration, specify a file name to store the calibration value.

To load an existing calibration file press the “Open File” button and browse and select the file.

4.2 Single measurement

Real-time measurements capture a sound with a measurement microphone connected to the UPV analyzer input 1, and calculate loudness and loudness level from the microphone signal. This is only possible if the microphone for capturing the sound has been calibrated. A single measurement is a real-time measurement performed once with a pre-defined acquisition time.

Select the “Single” tab and enter the SPL range that will be captured, the desired measurement duration, the field type and the block size of the samples that shall be processed. The possible block sizes of the samples are 48000, 24000, 4800, 2400 and 480. The default block size of 4800 is recommended. Minimum SPL range is 50 dBSPL, the maximum is 140 dBSPL. The maximum record length is 518 seconds.

To start the measurement press the “Start”-Button. An information box will pop up while the measurement is running in which pressing the “Cancel”-Button will cancel the measurement. When the measurement is completed, the SPL corresponding to a full-scale sine signal in the recording will be shown below the record length. It is calculated from the microphone calibration value and the used analyzer range. The displayed value can be used later on for offline re-evaluation of the recorded signal.

4.3 Offline evaluation

An offline evaluation calculates the specific loudness, percentile loudness and loudness level for the contents of an existing wave file (.wav). The sound pressure corresponding to a full-scale sine signal in the wav file has to be known. First the wav file has to be opened by clicking the “Open File” button and browsing to the desired file. The length and the sample rate of the sound will be shown automatically in the respective textboxes. The full scale value for the measurement must be selected in the numeric control. The maximum possible value is 160 dBSPL for a full-scale sine signal, and the minimum is -60 dBSPL. Also the field type and the block size of the samples for processing processes have to be selected. The possible block sizes depend on the sample rate of the file.

To start the measurement press the “Start” button.

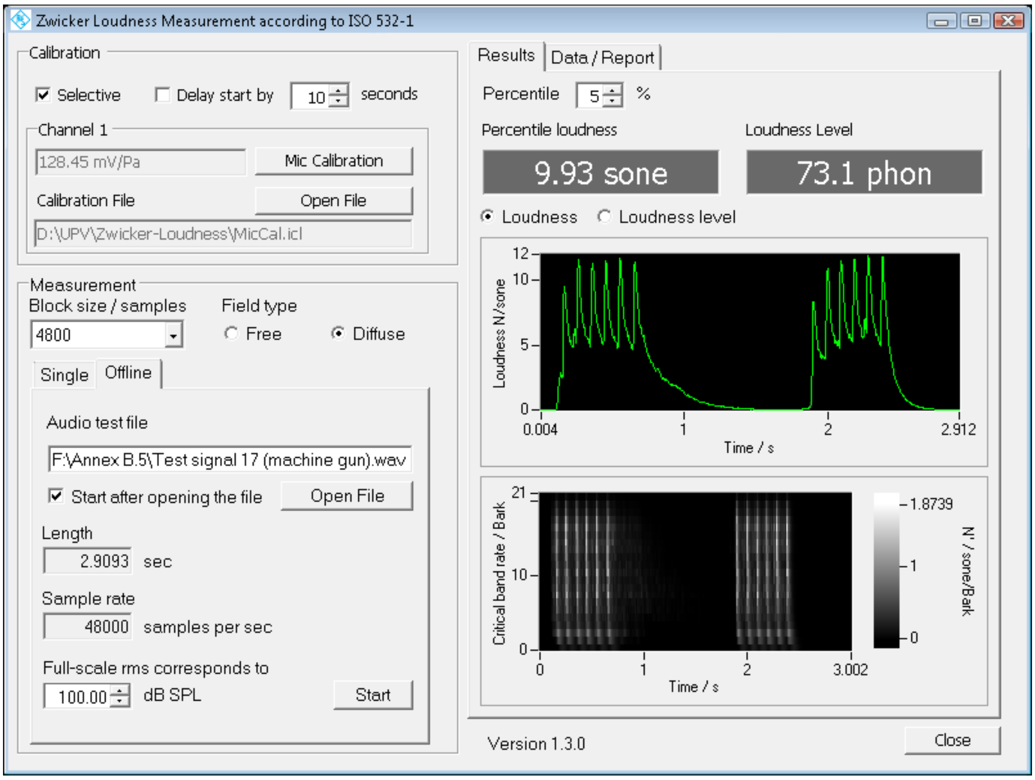

Fig. 4-2: Main window of the application program with sample result

4.4 Results

4.4.1 Percentile

The “Result” textbox will show the value for which the percentage of values selected in the numeric control are higher than this value if the checkbox for percentile is selected. To switch between loudness and volume level select the responding radio button and press the “Refresh” button.

For speech signals the 7% percentile is recommended, for all other signals the 5% percentile.

4.4.2 Diagram

After each measurement a little preview diagram of the results can be seen on the interface. To see the diagram in full size press the” Show Diagram” button. The application will minimize and the firmware with the diagram will be seen. In the UPV window the cursors can be used to further analyze the loudness versus time.

Whether a diagram shows the loudness or the volume level depends on which radio button is selected. To switch between loudness and volume level select the respective other radio button and press the “Refresh” button.

4.4.3 Saving data

For saving the loudness or loudness level data of each time step press the “Save Data” button. You can decide if you want to save it in a tabulator separated file (.txt) or in a comma separated file (.csv). The data always includes both loudness results and loudness level results. If the checkbox “Include specific loudness” is checked in the “Data / Report” tab, the values of specific loudness versus critical band rate and time are also included.

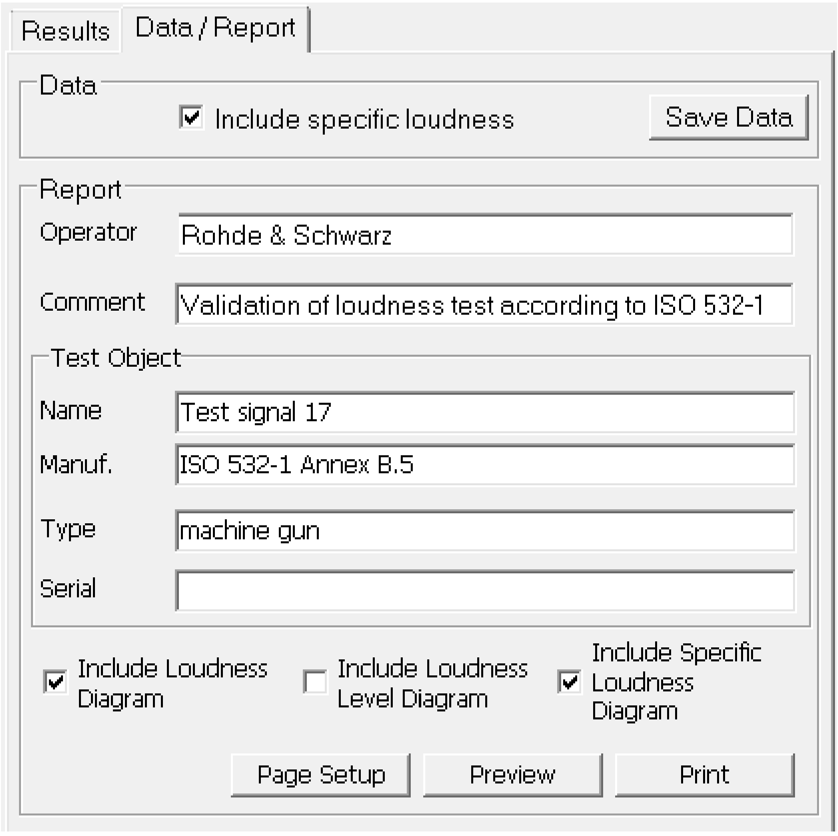

Fig. 4-3: Controls in the Data / Report tab

4.5 Report

4.5.1 Information

In the textboxes you can enter general information about measurements you have executed. This information will be shown at the beginning of the print page. You can append diagrams of the loudness and the volume level results to the report if you select one or both of the referring checkboxes.

4.5.2 Page setup

To change general settings of the print page like paper format, margins etc. press the “Page Setup” button.

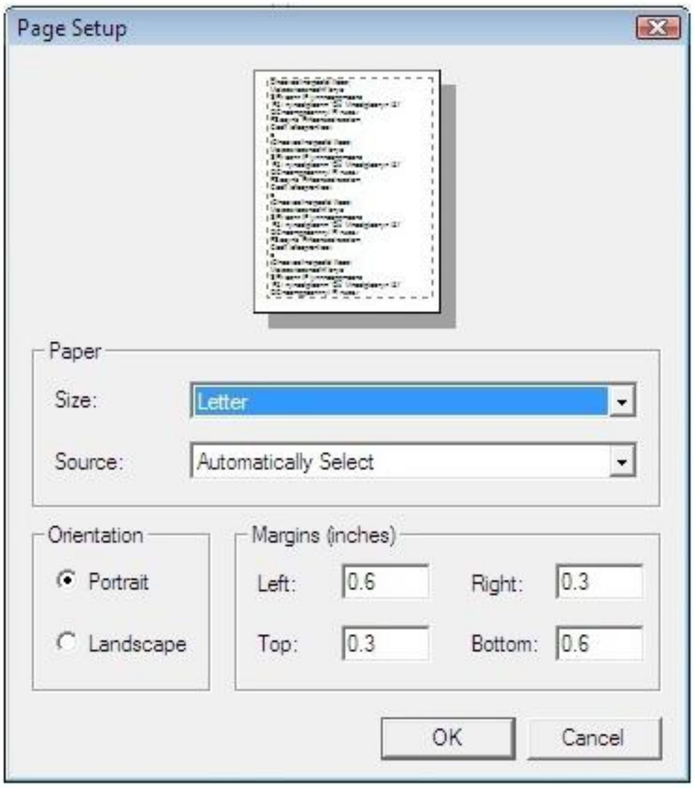

Fig. 4-4: Page setup dialog

4.5.3 Preview

To see a preview of the report pages with the settings you chose press the “Preview” button.

4.5.4 Print

The “Print” button opens the usual Windows print dialog where printer, print range and the number of pages can be selected.

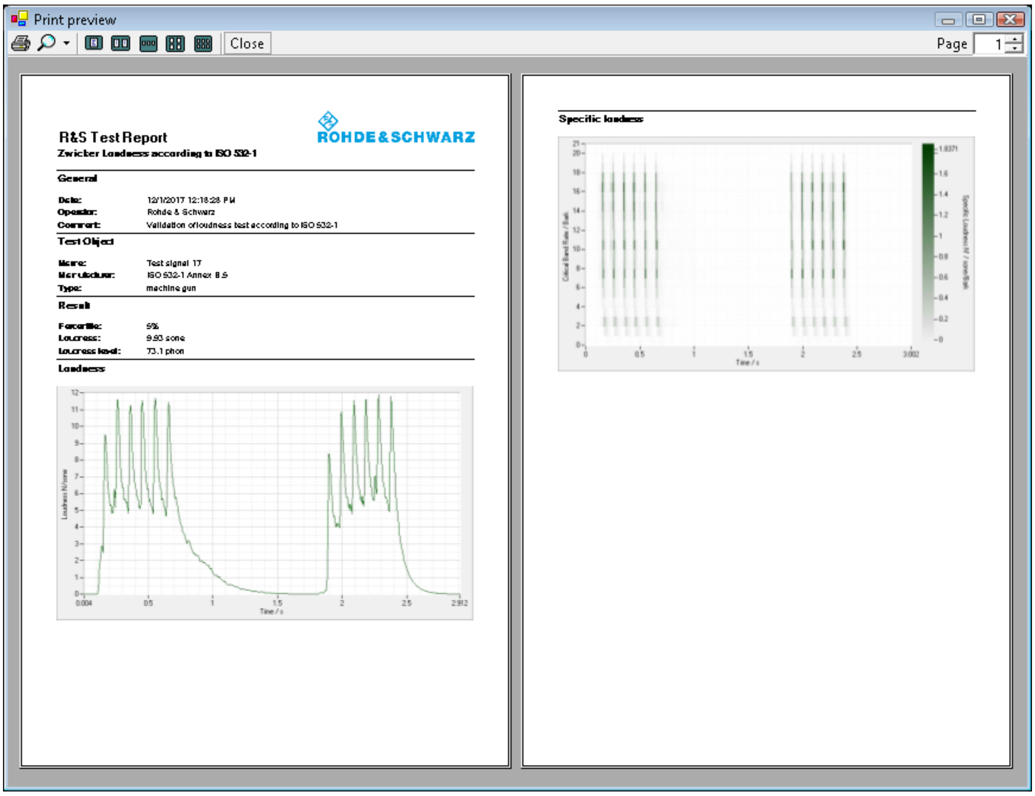

Fig. 4-5: Print preview

5 Literature

[1] **DIN** Berechnung des Lautstärkepegels und der Lautheit aus dem Geräuschspektrum - Verfahren nach E. Zwicker - Änderung 1: Berechnung der Lautheit zeitvarianter Geräusche; mit CD-ROM // DIN 45631/A1. - [s.l.] : DIN, 2010-03.

[2] **DIN** Berechnung des Lautstärkepegels und der Lautheit aus dem Geräuschspektrum; Verfahren nach E. Zwicker // DIN 45631. - [s.l.] : DIN, 1991-03.

[3] **E. Zwicker H. Fastl** Psychoacoustics - Facts and Models. - Berlin, Heidelberg, New York, London, Paris, Tokyo, Hong Kong, Barcelona : Springer-Verlag, 1990.

[4] **ISO** Acoustics — Methods for calculating loudness — Part 1: Zwicker method // ISO 532-1. - Geneva : ISO, 2017-06.

6 Ordering Information

Designation	Type	Order No.
Audio Analyzer	R&S®UPV	1146.2003.02
Universal Sequence Controller	R&S®UPV-K1	1401.7009.02

Rohde & Schwarz

The Rohde & Schwarz electronics group offers innovative solutions in the following business fields: test and measurement, broadcast and media, secure communications, cybersecurity, radiomonitoring and radiolocation. Founded more than 80 years ago, this independent company has an extensive sales and service network and is present in more than 70 countries.

The electronics group is among the world market leaders in its established business fields. The company is headquartered in Munich, Germany. It also has regional headquarters in Singapore and Columbia, Maryland, USA, to manage its operations in these regions.

Regional contact

Europe, Africa, Middle East
+49 89 4129 12345
customersupport@rohde-schwarz.com

North America
1 888 TEST RSA (1 888 837 87 72)
customer.support@rsa.rohde-schwarz.com

Latin America
+1 410 910 79 88
customersupport.la@rohde-schwarz.com

Asia Pacific
+65 65 13 04 88
customersupport.asia@rohde-schwarz.com

China
+86 800 810 82 28 |+86 400 650 58 96
customersupport.china@rohde-schwarz.com

Sustainable product design

- Environmental compatibility and eco-footprint
- Energy efficiency and low emissions
- Longevity and optimized total cost of ownership

Certified Quality Management
ISO 9001

Certified Environmental Management
ISO 14001

This and the supplied programs may only be used subject to the conditions of use set forth in the download area of the Rohde & Schwarz website.

Rohde & Schwarz GmbH & Co. KG
Mühldorfstraße 15 | 81671 Munich, Germany
Phone + 49 89 4129 - 0 | Fax + 49 89 4129 – 13777

www.rohde-schwarz.com

PAD-T-M: 3573.7380.02/02.05/EN/