Assessment of Changes in the Rise and Fall of the Audio Broadcasting Signal in the Transmission Channel

Oleg Popov
MTUCI
Moscow, Russia
olegp45@yandex.ru

Andrey Borisov
MTUCI
Moscow, Russia
a.borisov199@gmail.com

Tatiana Chernysheva
MTUCI
Moscow, Russia
krba2012@yandex.ru

Kirill Orlov
MTUCI
Moscow, Russia
kvororloff@yandex.ru

Valentin Abramov
MTUCI
Moscow, Russia
vabramov@mail.ru

Abstract— The change in the slope of the rise and fall of the audio broadcast signal is particularly noticeable to the listener. The slope of the rise determines the method of sound extraction, the recognizability of the musical signal, and the clarity of speech. Changes in the process of rise are noticeable to the listener in durations of just fractions of a millisecond. The process of the fall of the audio signal is mainly determined by the properties of the room in which it is recorded, such as reverberation. Changes in the process of fall affect the emotional color of the signal and are also very significant for the listener. Assessing the changes in the processes of rise and fall of the signal during its processing, coding, and transmission over the communication channel was difficult due to the lack of an objective evaluation method, which is proposed in the article.

I. Introduction

The problem of reducing distortion in analog and digital sound broadcast signals (SBS) arose simultaneously with the creation of systems for forming and transmitting broadcast signals. The basis for assessing quality has been and remains subjective statistical tests (SST), as the complexity of synthesizing the SBS and the lack of models for the perception of real-world sound signals make objective measurements difficult. Replacing SST with an objective psychoacoustic method of quality assessment is hampered by the lack of spectral analysis methods that would at least approach in accuracy the capabilities of the human peripheral auditory analyzer and the model of perception of a real sound signal [3]. Let us recall that existing models of perception are formed on the basis of studying the perception of harmonic and narrow-band noise signals. Increasingly, in problems of assessing transmission quality, researchers turn to the statistical characteristics of the sound signal [2, 4]. Forming objective estimates of the statistical parameters of signal characteristics for SBS is not a trivial task and required the development of a number of original algorithms. At the same time, algorithms were developed: for accurately estimating the power of a sampled real signal and the steepness of the process of increasing the amplitude of sound objects (ASO) of this signal [5], the rhythmicity of the appearance of sound objects of the signal [6], estimating the spectrum of the signal with accuracy, as close as possible to the accuracy of the auditory analyzer [7, 8]. In this case, sound objects in the signal are determined by sounds perceived as individual events (phonemes, notes). All of these algorithms use a complex representation of the signal. Attention is drawn to the fact that one of the most important parameters of sound signals, which significantly influence the quality of these signals and undergo noticeable changes in modern transmission channels, are the processes of increase and decrease in the amplitude of the signal over the duration of the existence of sound objects. Features of the functioning of the listed algorithms are given in works [8-10].

It is known that the elimination of the processes of increasing in the ASO, constituting about 15% of the duration of the speech signal, makes it completely illegible, and for a musical signal the ability to identify the instrument on which the melody was played is lost. In both cases, the semantic and emotional information of the sound signal is lost. Changes in the ASO decay process change the listener's sense of the spatial characteristics of the room in which the recording was made. The process of ASO decline is determined by the reverberation of the room and largely determines the emotional impact of the sound signal. The difficulty of objectively assessing changes in the rise and fall of ASO is determined by the fact that the spectra of the measured process and its assessment overlap each other. A convenient parameter for forming an objective assessment could be the analytical envelope of the audio signal. To form such an envelope, a complex representation of the audio signal must be formed. For this, an orthogonal signal must be formed from the wideband audio signal, with all of its spectral components rotated by 90°. Several variants of such a transformation have been proposed, including using the discrete Fourier transform (DFT) or its fast implementation (FFT), but the accuracy of forming the orthogonal signal was insufficient. We have managed to achieve the necessary accuracy of formation, with an error of no more than 10^{-5} instead of 10^{-3} from existing methods of analog representation of signals, optimizing and supplementing the window functions used in FFT. Using the complex representation of the signal, it became possible to form the analytical envelope of the audio signal, the low-frequency component of which well reflects the processes of rise and fall of the signal and allows its objective evaluation to be formed. Analysis of changes in this envelope in the channel showed the appearance of distortions at all stages of transmission, and therefore the need to correct these distortions. However, such correction turns out to be impossible when using existing automatic control algorithms. The developed original algorithm based on processing a precisely formed analytical envelope is given below. It allows you to compensate for distortions in the processes of increasing and decreasing ASO, and without violating the dynamic range. To implement the algorithms, software has been developed, some of which are given in [11, 12].

II. Algorithm for generating an orthogonal signal.

To represent a real signal in a complex form, it is necessary to generate a signal orthogonal to it, thereby implementing the Hilbert transform (HT). It is well known, that an ideal HT for physical signals is physically impossible. It is known that an ideal HT for signals is physically impossible, however, to solve various engineering problems it is possible to shift the errors of such conversion to an area where there is no signal energy, or to reduce these distortions to the threshold value of auditory perception.

In the case of the HT, the conjugate function f(t), associated with the original f(t), is expressed as:

$$\hat{f}(t) = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{f(\tau)}{t-\tau}d\tau = \frac{1}{\pi t}\cdot f(t), \qquad (1)$$

where function $\hat{f}(t)$ - conjugate to the original $f(t)$.

The Hilbert transform allows for a precise determination of modulation parameters such as the envelope and phase of a signal. In the case of assessing the rise and decay processes of a sound signal, it is possible to smooth the selected Hilbert envelope using a low-pass filter and applying a temporal window of approximately 10 ms [1]. The auditory sensitivity to changes in the envelope elements (attacks) is depicted in Fig. 1.

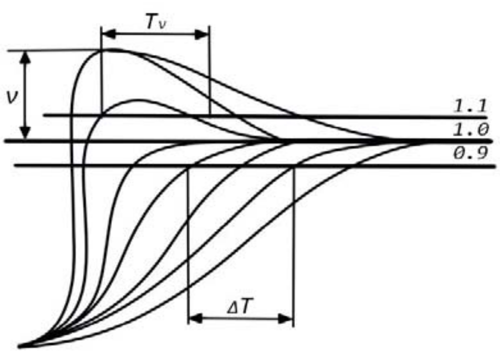

Fig. 1. Thresholds for the noticeability of changes in attack parameters of a sound object; explanation of the prolongation and emphasis of the attack.

Fig. 2 shows the visibility threshold as a function of attack length for various musical instruments, while Fig. 3 shows the audibility threshold of radiation as a function of its duration.

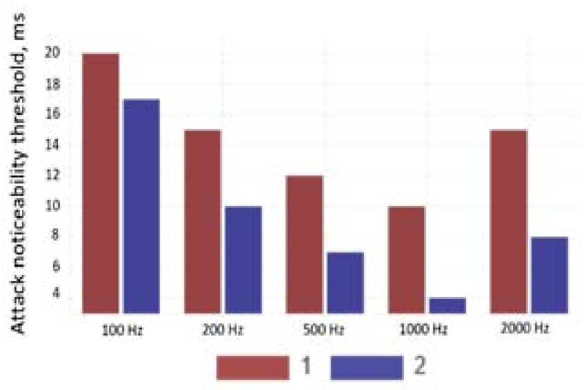

Fig. 2. Thresholds for the noticeability of changes in attack parameters of a sound object; delaying the attack of the piano, plucked (1) and bowed instruments (2).

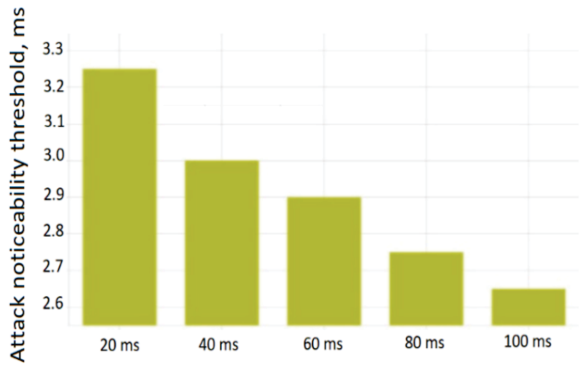

Fig. 3. Thresholds for the noticeability of changes in attack parameters of a sound object; threshold for the noticeability of the attack accent as a function of its duration.

III. Accuracy of orthogonal signal generation based on FFT.

Synthesizing an orthogonal signal using the Discrete Fourier Transform (DFT) is possible. There are two methods for Hilbert transformation of wideband signals. The first method involves amplitude modulation and phase rotation of the signal. However, the accuracy in this case is not very high (up to 10^{-3}).

The second method involves using the orthogonal Hilbert transform directly in the frequency domain with a 90° phase shift of the DFT coefficients. When using the Fast Fourier Transform (FFT), this is a simple and fast algorithm [3], based on the formula:

$$\hat{X}_k = X_k * H_k, H_k = \begin{cases} -j,\ k = \left(0, \frac{N}{2} - 1\right) \\ +j,\ k = \left(0, \frac{N}{2} - 1\right) \end{cases}, \qquad (2)$$

where: X_k - DFT coefficients of the original signal, $\hat{X}_k$ - coefficients of the orthogonal signal, H_k sequence - can be interpreted as a discrete representation of the jsign (ω) function. The accuracy of synthesizing an orthogonal signal is increased when individual groups of DFT coefficients representing spectral components do not overlap. This can be achieved by using window functions with minimal side lobes during transformation. Extending the main lobe can be compensated by increasing the length of the sample [2, 3].

Leakage can distort amplitudes and mask weak signals, making them harder to detect. Using different window functions can reduce the level of side lobes but leads to a compromise between the width of the main lobe and frequency resolution.

In Fig. 4, it can be seen that the Nuttall window has a lower signal-to-noise ratio compared to other windows, but only up to coefficient 3. For example, in the fourth interval, the noise level for the Nuttall window was 85.7 dB (A) and remained almost constant. By adjusting the windows up to four coefficients, this error can be reduced to 10^{-5}, allowing the use of HT algorithms. Such accuracy is achieved through Fourier transformation and the modified Nuttall window.

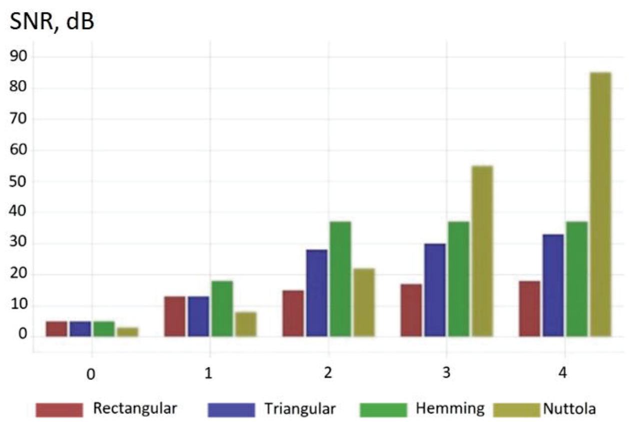

Fig. 4. Dependence of the signal-to-noise ratio (SNR) for different window functions.

The algorithm for synthesizing a high-precision orthogonal Hilbert transform signal is presented in Fig. 5. It includes the following operations:

• inputting the signal and applying a window function;
• phase-shifting all spectral components by 90 degrees;
• adding the previous sample and compensating for window irregularities.

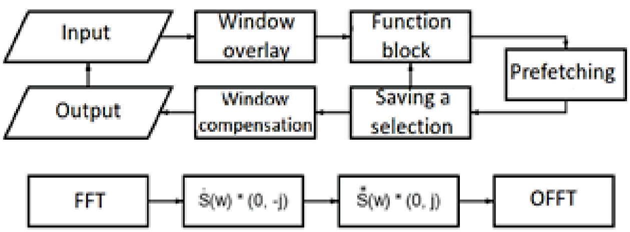

Fig.5. Scheme of the orthogonal signal synthesis algorithm

The error in generating an orthogonal signal using the DFT transformation is largely determined by the removal of the test frequency from the edge of the frequency range. The issue is that the zero coefficient and the coefficient corresponding to the Nyquist frequency do not have an imaginary part, making it impossible to change their phase. This introduces an error in the neighboring coefficients, which can be reduced by increasing the length of the sample and filtering out frequencies beyond the audible frequency range. The signal-to-noise ratio for this implementation is shown in Fig. 6 [2, 5].

According to the results presented in Fig. 6, the accuracy of synthesizing frequencies corresponding to integer coefficients of the DFT is significantly higher, as the entire signal energy is estimated by a single coefficient.

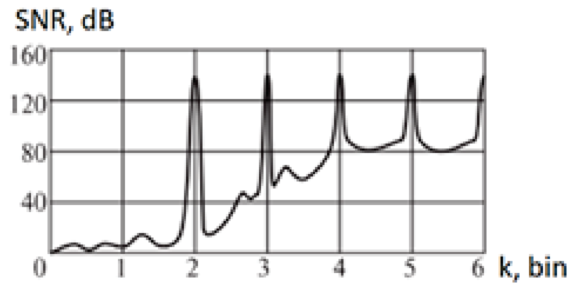

Fig. 6. Signal-to-Noise Ratio in Orthogonal Signal Synthesis Using FFT.

The algorithm includes the operation of compensating for the irregularity of the Nuttoll window at 50% overlap [6]. The use of the FFT basis in the implementation of the Hilbert transform provides for an increase in the sample length with the possibility of discarding part of the results, with an error in the synthesis of the orthogonal signal dsit 10^{-5}. This accuracy is sufficient for efficient analysis of the nature of the change of the analytical envelope during processing, compact representation and transmission of the audio signal Broadcasting. Software was developed to implement the algorithm [9]. Fig. 7 shows an example of the formation of an analytical envelope for a segment of the real sound of a symphony orchestra.

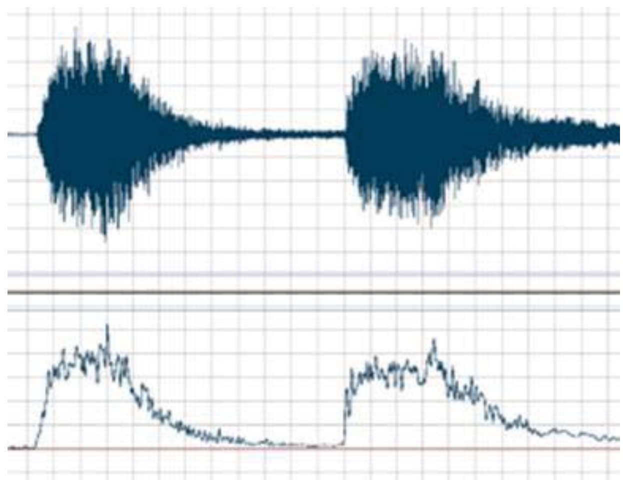

Fig. 7. Waveform of the signal segment and its analytical envelope.

IV. ALGORITHM FOR STATISTICAL EVALUATION OF THE PROCESSES OF INCREASE AND DECREASE IN THE AMPLITUDE OF SOUND OBJECTS IN AN AUDIO SIGNAL

Distortions in the rise of the amplitude envelope (rate of level change more than 200 dB/s) and fall (rate of change less than 200 dB/s) are evaluated based on changes in the slope values of the low-frequency component of the analytic envelope of the signal, which unambiguously corresponds to the processes of real signal change.

The algorithm includes the following operations on the input:

• calculation of the analytic envelopes of the signals;

• low-pass filtering of the analytic envelopes;

• segmentation of the low-frequency envelope into intervals corresponding to the rise and fall sections of the signal.

• calculation of the slope of the analytic envelope in the rise and fall sections;

• calculation of the relative change in slope between the analytic envelope of the original and distorted signals;

• calculation of the mean value, standard deviation, median slope values in the rise and fall sections, as well as the integral absolute deviation of statistics (IADS) for the slope distributions of the envelope in non-stationary signal intervals:

$$S_A = \frac{\left|\tilde{A}_{i+k} - \tilde{A}_i\right|}{k},$$

where $\tilde{A}_i$ is the values of the low-frequency envelope of the signal.

V. STUDY OF SPECIFIC DISTORTIONS IN THE TRANSMISSION CHANNEL.

To conduct research, a database of sound signals from various genres was created. Preliminary analysis showed that the symphony orchestra has the largest dynamic range and variations, and its sound was used in further investigations.

One of the most common procedures that should not introduce distortions in the signal is the modification of the spectrum width of the transmitted signal. As known, the frequency range of a signal in a studio is 20-20000 Hz, for broadcasting, bandwidths up to 15000 Hz are used, for signals with higher quality class (HQ), up to 10000 Hz, and for signals with sound class 1 and 2 (1st class, 2nd class), up to 6300 Hz. The normalized steepness of the envelope's rise (attack) and fall (decay) relative to the maximum value is shown in Fig. 8.

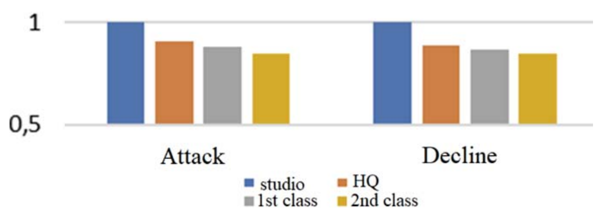

Fig. 8. Normalized steepness of attacks and declines after signal filtering according to different quality classes.

Thus, even a routine, ubiquitous signal filtering operation leads to a change in the steepness of attacks and signal drops, distortion that is especially noticeable when listening to a real signal. Unfortunately, transmission quality measurements are usually made on stationary signals, and only the latest quality assessment recommendations attempt to use real signals in a limited set and short duration [7, 8].

They contribute to the distortion of the processes of rise and fall of the signal and other trivial operations on the sound signal, which are performed in the studio and distribution channels throughout the country. One of these operations is audio processor signal level processing. In the studio, they strive to increase the volume (power) of the transmitted signal within a limited level. As a rule, this is accompanied by a compression of the dynamic range. For example, when two speakers are running on the same microphone, a male and a female, a compression of ap-proximately 18 dB is used.

Fig. 9 shows the results of measuring the slope of the edges for compression of the dynamic range of the audio signal using a dynamic audio processor at 10 and 20 db.

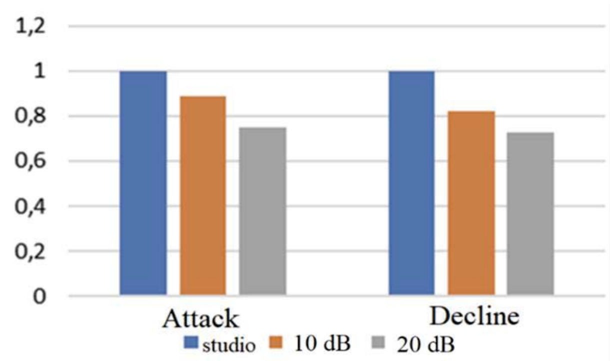

Fig. 9. Dependence of rise and fall steepness on the value of compression of the dynamic range of the signal.

As we can see, this ubiquitous operation also distorts the steepness of the real signal. In order to avoid intra-signal signal modulation, multiband dynamic processors are used in practice. In this case, the signal is first divided into bands, then adjusted, in each band, and restored again. The results of the study of changes in rise and fall for a multiband audio processor with a decrease in the dynamic range of the signal by 10 dB are shown in Fig. 10.

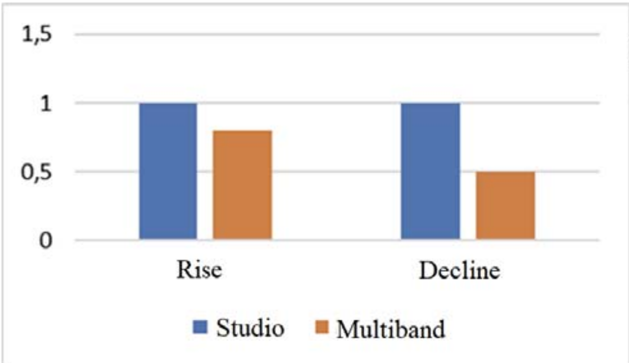

Fig. 10. Change in rise and fall slope in processing in a multiband dynamic processor.

The slope changes in a multiband audio processor are much greater than for a single-band audio processor, be-cause the rise and fall of each sub band signal is now determined not only by the properties of the signal, but also by the properties of the filter.

Another commonly used operation in the transmission channel is the change (conversion) of the sample rate, which is performed simultaneously with the filtering of the signal. The results of the study of the effect of sample rate converters on the slope of rise and fall of the signal are shown in Fig. 11.

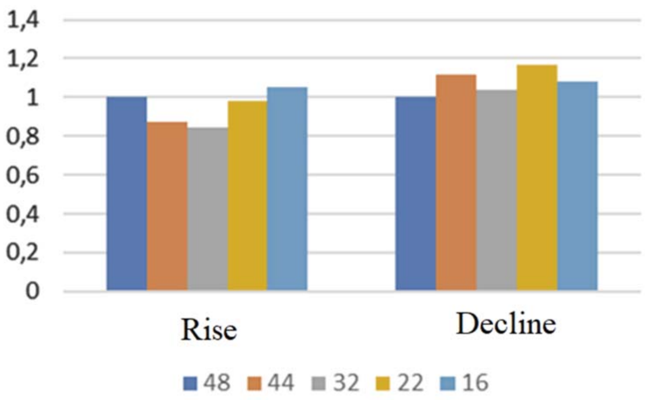

Fig. 11. Change in steepness during sample rate conversion.

The change in steepness when changing the sample rate is less natural than with other types of processing, since it depends on the spectrum of the signal and the relationship between the spectral components of the signal and the sample rate. The new values found for the new time positions may be closer to the true peak value of the original signal and have a correspondingly larger amplitude. Given the harmonic structure of the music signal, such an increase in amplitude can exceed 20 dB, which is taken into account by modern rules for regulating peak signal levels in the studio. Therefore, after changing the sample rate, the slope of the signal can both increase and decrease.

A more unambiguous relationship is due to changes in the steepness of rise and fall of the ASO systems with a compact representation in the frequency domain, for example, when encoding sound in the MP3 format. Fig. 12. The results of measuring the steepness of rise and fall when using different transmission rates are given.

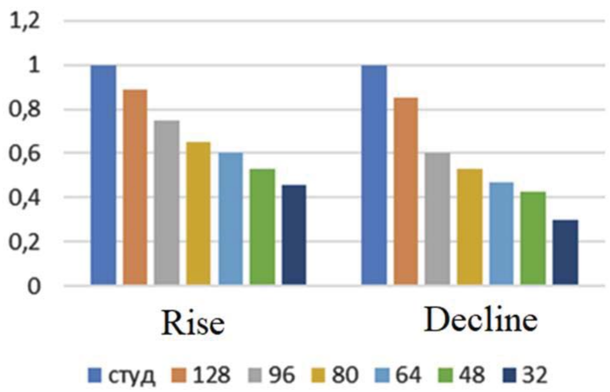

Fig. 12. Dependence of the rise and fall of the signal on the transmission rate in the compact representation in the MP3 format.

The transition to the frequency domain is carried out by a set of about 30 filters (depending on the embodiment) with a bandwidth close to the "critical" hearing strips. After this transformation, the rise and fall process is determined mainly by the properties of the filter. This is especially evident when several codecs are used in series in the transmission channel. In this case, the "attacks" and fallbacks become the same for all sound objects (sounds).

Studies were carried out on the maximum and mini-mum deviations of the steepness of rise and fall, as well as the visibility of these changes to the listener. The results of the studies are shown in Fig13.

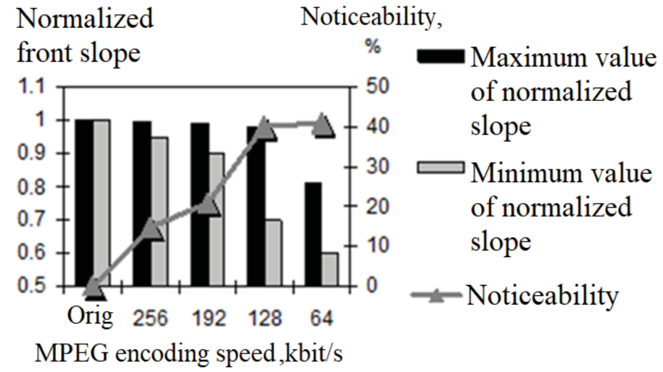

Fig. 13. Dependence of the steepness of rise and fall on the speed of transmission in MP3 format and the visibility of these changes for the listener.

Distortion of the steepness of the ASO is one of the noticeable distortions of the sound signal in modern transmission channels using a compact representation of the signal. It is noted that the distortion of the increase of the ASO of the signal makes it impossible to determine the process of sound production, and, consequently, the identification of the instrument. At the same time, not only the recognizability of the musical signal decreases, but also the intelligibility. Unfortunately, the existing audio processing algorithms do not allow to effectively correct these distortions, but rather add new ones, Fig. 9, 10 [9, 10].

One of the few algorithms that allows you to restore the signal to some extent is the method of processing the analytical envelope without compressing dynamic range, proposed by the authors [12]. The work proposes to regulate the analytical envelope using a device with the characteristic shown in Fig. 14.

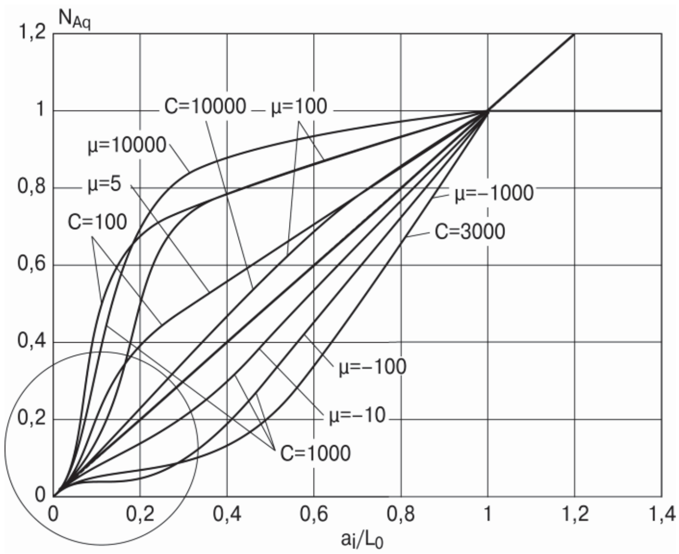

Fig. 14. Amplitude response for analytic envelope control.

The characteristic adaptively changes its shape according to changes in the relative average power of the input signal, which allows you to adjust the power and therefore the volume of the signal. It is possible to avoid the appearance of nonlinear distortions by dividing the envelope into low-frequency and high-frequency parts, the low-frequency component changes with the speed of sound objects, up to 5-7 per second. In practice, this is a constant component that is not distorted during regulation. The high-frequency component of the analytical envelope changes in proportion to the change in the low-frequency one, which also does not lead to the enrichment of the signal spectrum. In order to implement the algorithm in the mode of compensating for slope distortion in the channels of the sound broadcasting signal, software was used [9], with the help of which the possibility of compensating for the distortion of steepness in the transmission channels with the elimination of redundancy was studied. Compensation was carried out in two modes:

- in the pre-emphasis mode, i.e., in front of the distorting channel, for example, at the output of the programming path;
- in the correction mode, which involves the restoration of the signal at the receiving point. The results of the study are shown in Fig. 15.

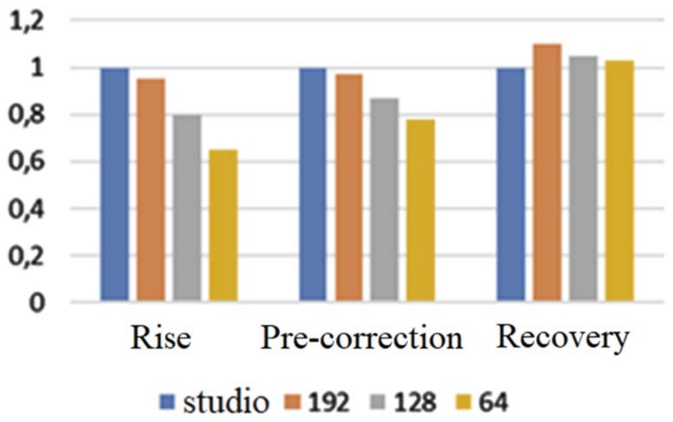

Fig. 15. Compensation of the signal build-up process with the help of preliminary and restorative correction

According to the results of the experiment, with the help of preliminary correction, it is possible to reduce the amount of distortion in the process of increasing the sound broadcast signal. The advantage of pre-correction is the installation of the corrector at the point of the central transmitter. Restorative correction restores the coolness of the signal, but in this case, a corrector must be installed at each receiving point.

VI. Conclusion

It is shown that by now the distortions, but rather the signal changes in the audio broadcast signal transmission channel are very large and are not covered by the existing methods of measuring objective signal parameters. This applies to almost all algorithms for processing and compact representation of SBS, therefore, statistical data on changes in signal parameters are increasingly used to assess transmission quality.

Attention is drawn to the fact that one of the most important parameters of sound signals, which significantly influence the quality of these signals and undergo noticeable changes in modern transmission channels, are the processes of increase and decrease in the amplitude of the signal during the duration of the existence of sound signals objects.

It was found that the use of statistics on the steepness of the rise and fall of the amplitude of sound objects makes it possible to carry out an objective assessment of distortions and evaluate the effectiveness of original methods of compensating them.

It has been shown that changes or distortions in the processes of increase and decrease in the amplitude of sound objects in the signal arise already at the stage of filtering, analog-to-digital conversion, signal level processing and continue at all stages of transmission of an audio broadcast signal through a transmission or storage channel and, consequently, Unfortunately, they are not always taken into account by the developers of transmission systems.

The study showed the insufficiency of the existing metrological support for transmission and storage channels of the audio broadcasting signal (SBS). Quality assessment is still made by changing the shape of a low-component, harmonic stationary signal, while almost all modern channels do not preserve this shape.

It is shown that the only method for assessing quality remains subjective statistical tests, which are not applicable when it is necessary to quickly assess the quality of transmission and its changes, for example, due to new coding methods or audio processing. There is a need to study and identify signal parameters, changes in which determine the listener’s assessment of the transmission quality.

It was found that existing methods for compensating for distortions do not allow compensating for distortions in the processes of increase and decrease in the amplitude of sound objects of a signal when it is transmitted over a communication channel.

The results of the work done were:

- an algorithm for complex presentation of the audio signal has been developed;
- an algorithm has been developed for generating an estimate of the increase and decrease in the amplitude of sound objects of the analytical signal envelope;
- distortions in the processes of increase and decrease in the amplitude of sound objects of a real sound signal in specific transmission channels have been studied;
- a method has been developed for preliminary and restorative correction of distortions in the processes of increase and decrease in the amplitude of sound objects of the signal envelope;
- software was developed to implement the proposed algorithms;
- measurements were carried out on a real signal, which confirmed the possibility of compensating for distortions in the processes of increase and decrease in the amplitude of sound objects in the envelope of this signal.

Using the results of this work makes it possible to improve the quality of broadcast signals and information programs, as well as to better regulate audio broadcasting equipment and thereby increase the popularity ratings of broadcast stations, as well as increase the economic efficiency of these stations.

Literature

[1] Y. A. Kovalgin, Digital Radio Broadcasting: Systems and Technologies. Hotline - Telecom. 2021, 580 p.

[2] Ntalampiras, "Automatic analysis of audio streams in the concept drift environment," 2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP), Vietrisul Mare, Italy, 2016, pp. 1-6, doi: 10.1109/MLSP.2016.7738905. keywords: {Hidden Markov models; Feature extraction; Dictionaries; Adaptation models; Computational modeling; Pattern recognition; Signal processing; Computational auditory scene analysis; Concept drift; online adaptation and modelling}.

[3] V. A. Abramov, O. B. Popov; T. V. Chernysheva; V. O. Peruanskiy Increasing the Accuracy of Sound Signal Spectral Estimation According to the Properties of Hearing Analyzer Intelligent Technologies and Electronic Devices in Vehicle and Road Transport Complex (TIRVED) DOI: 10.1109/TIRVED53476.2021 11-12 Nov. 202.1

[4] Recommendation ITU-R BS.1387-2 (05/2023) BS Series: Broadcasting service (sound) Method for objective meas-urements of perceived audio quality.

[5] V. A. Abramov, O. B. Popov, V. G. Taktakishvili, A. A. Ovchinnikov «Method and device for measuring power and steepness of increase in sections of nonstationary of acoustic signals», RU Patent 2731339, Jan. 9, 2020.

[6] V. A. Abramov, O. B. Popov, I. V. Vlasyuk, A. V. Balobanov “Method and device for measuring rhythmic frequencies, power and duration of drops in the non-stationary areas of acoustic signals”, RU Patent 2 773 261, Jun. 1, 2022.

[7] V. A. Abramov, O.B. Popov, I. V. Vlasyuk, A. V. Balobanov, «Method and apparatus for measuring the spectrum of informational acoustic signals with distortion compensation», RU Patent 2756 934, Oct. 7, 2021.

[8] V. A. Abramov, O. B. Popov, T. V. Chernysheva, A. A. Borisov, “Method and apparatus for high-precision measurement of the spectrum of informational acoustic signals”, RU Patent 2808156, Nov. 24, 2023.

[9] V. A. Abramov, O. B. Popov “Program for high-precision spectral analysis of the sound signal. Software "DKP-Spektr". Certificate of state registration of software No. 2022 2022616423. Date of registration: April 19, 2022.

[10] V. A. Abramov, T. V. Chernysheva, O. B. Popov, V. O. Peruanskiy and P. G. Kuznetsov, "Frequency Analysis of Sound Broadcasting Signal Parameters," 2022 Systems of Signals Generating and Processing in the Field of on Board Communications, Moscow, Russian Federation, 2022, pp. 1-7, doi: 10.1109/IEEECONF53456.2022.9744387.

[11] V. A. Abramov, O. B. Popov, V.G. Taktakishvili, T.V. Chernysheva, A.A.Ovchinnikov “Complex representation of signals by the "Complex" software. Version 2019”. Certificate of state registration of software No. 2019667523. Date of state registration: December 24, 2019.

[12] V. A. Abramov, O. B. Popov, V. G. Orlov, “Method and apparatus for companding audio broadcasting signals”, RU Patent 2 691 122, Jun. 11, 2019.