Forming the Frequency Response Assessment of the Audio Transmission Channel Based on the Real Signal

Oleg Borisovich Popov
MTUCI
Moscow, Russia
olegp45@yandex.ru

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

Tatiana Vasilyevna Chernysheva
MTUCI
Moscow, Russia
krba2012@yandex.ru

Kirill Vaielevich Orlov
MTUCI
Moscow, Russia
kvororloff@yandex.ru

Valentin Aleksandrovich Abramov
MTUCI
Moscow, Russia
vabramov@mail.ru

***Abstract*—** The frequency response (FR) of the audio broadcasting channel is formed, as a rule, using a test sinusoidal signal. In the process of transmission via modern adaptive channels, the real sound broadcasting signal (SBS) is subjected to compact representation, audio processing and radiation using acoustic systems. All these procedures significantly change the FR of the signal, but retain subjectively comparatively high transmission quality. The original methods of spectral analysis developed by the authors made it possible to bring their accuracy and resolution closer to the previously unattainable capabilities of the human auditory analyzer. This, in turn, makes it possible to provide the ability to form and control the FR directly in the process of transmitting the audio program for all devices included in the channel.

I. Improving the Accuracy of Spectrum Estimation

The most convenient method for forming spectral assessments remains the discrete Fourier transform (DFT), or its computationally efficient implementation, the Fast Fourier Transform (FFT). The FFT transforms a signal, discretely described in the time domain, into a set of coefficients in the frequency domain. Unfortunately, the necessary accuracy for forming assessments of the amplitudes and phases of spectral components is only provided by this transformation for frequencies that coincide with the frequencies of the FFT coefficients, known as bins. It has been proposed [1,2] to enhance the accuracy of frequency assessments located between the bins by repeatedly transposing the original signal spectrum through multiplication of its complex representation by a complex representation of the frequency within the bin fractions:

$$S\gamma(n) = S(n)\cdot \exp(j\cdot 2\pi + \gamma). \quad (1)$$

The maximum amplitude Sγ(n) corresponds to the coefficient with the closest approximation of the FFT coefficient parameters to the true values of the spectral component parameters. An orthogonal signal generated using the FFT is used to form the complex description of the input signal.

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

To assess the accuracy of the transformation at the first and last coefficients, an analysis of the window function parameters in the frequency domain was conducted. For these coefficients, the ratio presented below was calculated.

$$R_k = \frac{\sum_{i=1}^{N/2-1} E_i}{\sum_{i=k+1}^{N/2-1} E_i} \quad (3)$$

Where Ei is the energy of the i-th coefficient, N s the coefficient number to which the energy of all coefficients is related. This ratio defines the maximum theoretical error in synthesizing the orthogonal signal for the harmonic signal represented by the i-th coefficient, which is determined by the properties of the window.

Fig. 1 shows the graphs of R_k or four windows: Nuttall, Hamming, triangular, and rectangular. Based on the analysis results, the Nuttall window was chosen for forming the orthogonal signal. Fig. 2 presents the signal-to-noise ratio (SNR) of the orthogonal signal as a function of the frequency of the oscillation being transformed for six values of the parameter Fd/N (bin), which characterizes the frequency band represented by a single coefficient. The error in forming the orthogonal signal does not exceed 10^{-5} when using the Nuttall window function in the synthesis process and a specially developed compensating function after the DFT, with a duration of the discrete sequence of at least 4000 points [1,2,3,4].

$$w(n) = a_0 - a_1 \cdot \cos\left(\frac{2 \cdot \pi \cdot n}{N-1}\right) + a_2 \cdot \cos\left(\frac{4 \cdot \pi \cdot n}{N-1}\right) - a_3 \cdot \cos\left(\frac{6 \cdot \pi \cdot n}{N-1}\right), \quad (4)$$
$$a_0 = 0.355768,\ a_1 = 0.487396,\ a_2 = 0.144232,\ a_3 = 0.012604.$$

To compensate for the non-uniformity of the transfer coefficient of this window after the DFT with 50% overlap, an additional window $1/W^2(n)$ is used.

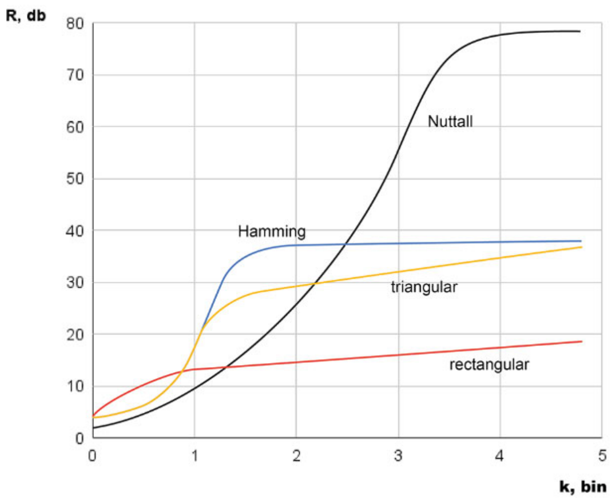

Fig. 1. Ratio R_k for four window functions

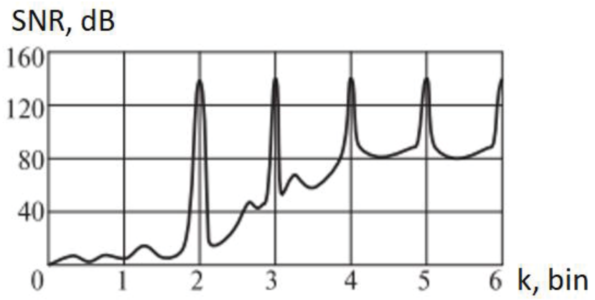

Fig. 2. Error in synthesizing the orthogonal signal for a harmonic oscillation

The developed algorithm allows for approximating the accuracy of spectral component assessments to the accuracy of the human auditory analyzer, with a frequency calculation accuracy of 1.5 Hz in the range below 500 Hz and 1.5% of the absolute frequency value in the range above 500 Hz; amplitude calculation accuracy of 0.4 dB; phase calculation accuracy of 8 degrees (according to SST results); resolution of no worse than 60 Hz; and a signal duration of no more than 8 ms (according to SST results).

An unavoidable drawback of the Discrete Fourier Transform (DFT) is its insufficient resolution, meaning the ability to distinguish between two closely spaced spectral components. As is known, each DFT coefficient evaluates the energy of the signal in the band defined by the formula:

$$df = \frac{B \cdot f_D}{N}, \quad (5)$$

where:

B is the window expansion coefficient defined by the window function;

f_D is the sampling frequency;

N is the length of the discrete data sample.

The Discrete Cosine Transform (DCT) is widely used, which has a band for evaluating the signal that is half as wide, defined by the formula:

$$df = \frac{B \cdot f_D}{2N} \quad (6)$$

The DCT of the data array L(m), where m = 0, 1, …, N–1, is defined as follows:

$$L_X(0) = \frac{1}{\sqrt{N}} \cdot \sum_{m=0}^{N-1} X(m)$$
$$L_X(k) = \sqrt{\frac{2}{N}} \cdot \sum_{m=0}^{N-1} X(m) \cos\frac{(2m+1)k\pi}{2N}, \quad (7)$$
$$k = 1, 2, \ldots, N-1$$

A disadvantage of the DCT is the wide spectrum of side lobes and the lack of the ability to represent the signal in a complex form. Both of these drawbacks have been overcome by the authors of the article, allowing for the formation of accurate assessments of instantaneous spectra over short time intervals.

II. Investigation of Transfer Coefficients of Adaptive Paths in the Audio Transmission Channel

There is a difference between the transfer coefficients at different frequencies for a harmonic stationary sinusoidal signal, used to assess the channel's transmission quality, and the actual multi-component sound broadcasting signal (SBS). It was precisely to identify these deviations that we developed an algorithm for spectral analysis with increased accuracy [1,4] and software for its implementation. To confirm the possibility of forming a frequency response (FR) based on a real signal, an experiment was conducted to identify the known distortions, shown in Fig. 3, introduced into the AS. One of the difficulties in spectral analysis of a real AS is the need to use sufficiently long signals for collecting statistics. In this example, Liszt's Rhapsody No. 2 was used as the signal. The choice of the national anthem as the signal is justified by the fact that many radio stations use the transmission of the anthem at the beginning of the broadcast, and checking the quality of its transmission allows for assessing the current state of the channel. Fig. 3 shows the waveform of the signal, against which the FR of the filter used for processing is shown.

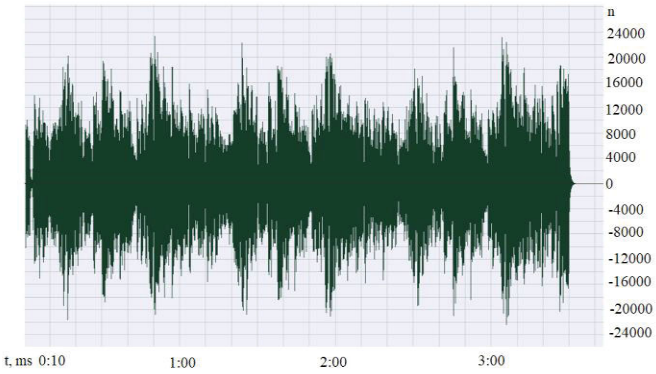

Fig. 3.1 Oscillogram of the signal of Liszt's Rhapsody No. 2

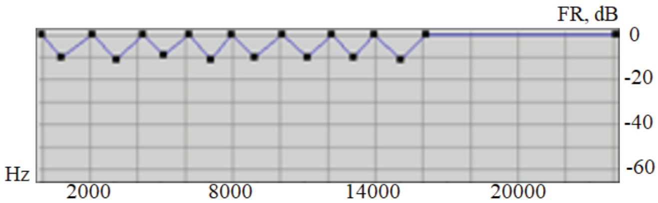

Fig. 3.2 FR processing filter

Fig. 4 shows the results of forming the FR based on a real signal against a template that normalizes the channel quality with a bandwidth up to 10 kHz. As can be seen from the Fig 4, the introduced distortions are confidently identified when forming the FR based on a real signal with a duration of 2.8 s.

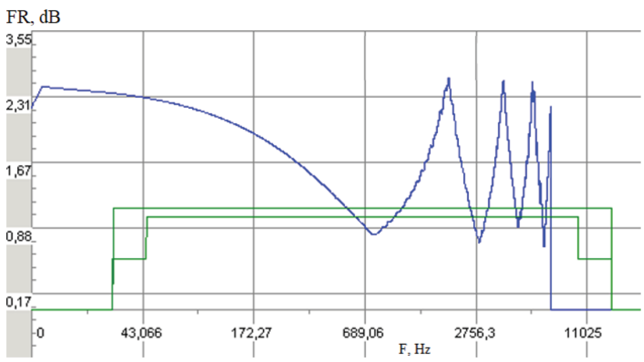

Fig. 4. FR response of the deformed signal of Liszt's Rhapsody No. 2

The results of assessing the frequency response of the transmission channel using the signal of Liszt's Rhapsody No. 2 are presented in Fig. 5.

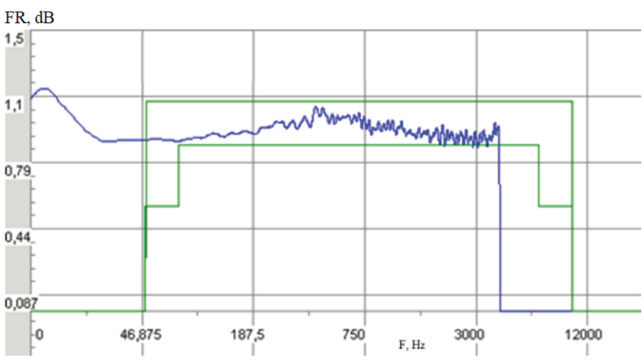

Fig.5. FR of a real channel generated from the signal of Liszt's Rhapsody No. 2

To date, virtually all audio transmission channels, both analog and digital, have become adaptive and alter the signal properties according to the capabilities of the transmission path. Audio processing of the SBS is used intensively. Moreover, the nature of this processing significantly depends on the signal properties. When transmitting a test sinusoidal signal, the processing is typically disabled, and the device switches to a linear amplitude characteristic. Additionally, many recommendations for assessing transmission quality suggest disabling all "additional" signal processing devices when evaluating the FR of the channel.

Such devices include companding systems, automatic level controllers, audio processors, noise suppressors, and so on. As a result, distortions that occur during the transmission of real signals often remain undetected during control measurements. An example of such distortions in the FR of the Optimod audio processor from Orban, identified on a real broadcast signal, is shown in Fig. 6.

The audio processor performs multi-band signal processing, and these bands, which are not detected by test signals, become apparent during the transmission of real signals.

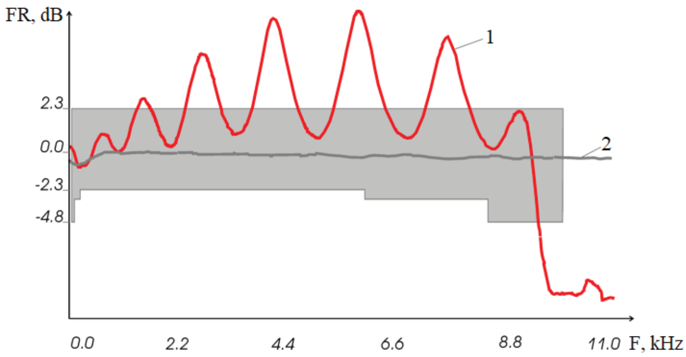

Fig.6. FR of the Optimod audio processor, Orban, formed from a real signal (1) and ARGO (2)

The signal was formed during the transmission of a real program with a duration of about 30 minutes, which affected the smoothing of the graph. In the same graph, the FR of the multi-band automatic gain control (AGC) with Hilbert envelope, developed by the authors of the article, is shown in light gray. For ease of operation, audio processors with a large number of preset modes are used, the influence of which on the signal is not always accurately assessed by sound engineers. For a long time, listeners noted the "velvety" sound of the radio station, which did not always correspond to the nature of the transmitted program. An inspection of the program formation path revealed that the T.C. Electronics DBMax audio processor was used in the studio, which emphasized the low-frequency components on the real signal, a characteristic that was not detected with test harmonic signals. Fig. 7 shows the FR of the program formation path, which includes the specified audio processor, and formed based on the real broadcast signal.

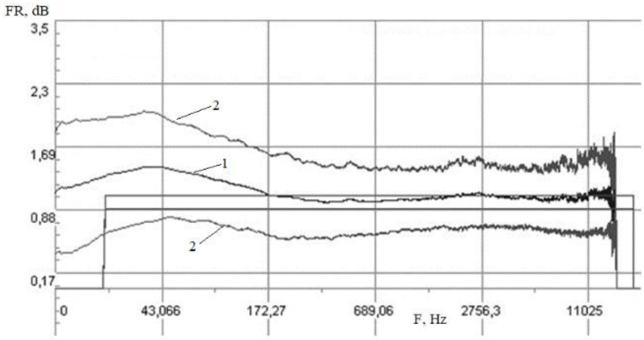

Fig. 7. FR of the audio processor T.C. Electronics DBMax in the on-air studio (1-average value FR, 2- dispersion of values FR)

Moreover, similar settings were established on the audio processors in both the main and backup studios, as shown in Fig. 8.

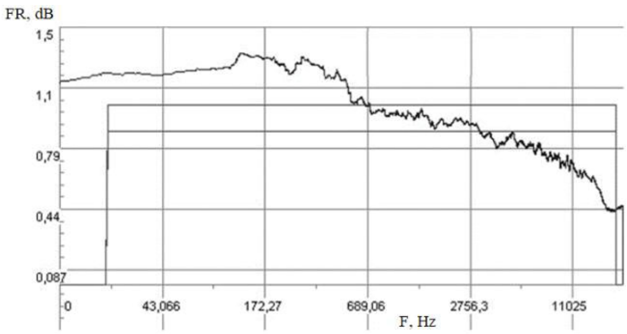

Fig. 8. FR of the audio processor T.C. Electronics DBMax in hardware backup studio

To get rid of the "velvety" sound, it was necessary to change the audio processor.

In the process of forming the SBS for transmission, many methods of processing and introducing sound effects are used, one of the most common being the correction or deformation of the frequency characteristic of the signal [4, 7, 8]. Such a trivial operation also needs to be controlled using a real signal. Fig. 9 shows the frequency characteristic of the equalizer formed in the Adobe Audition editor. Control using a sinusoidal signal confirms the correspondence of its characteristics to the interface.

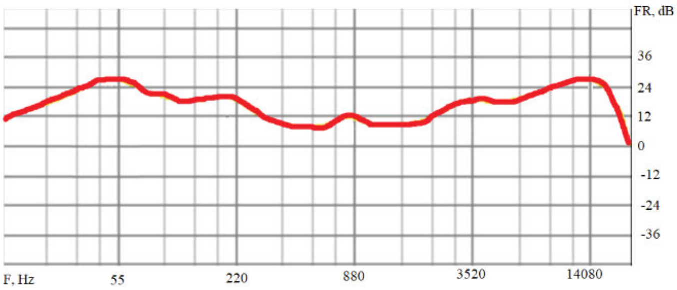

Fig. 9. Equalizer FR

When checking the operation of the equalizer on a real sound signal, the characteristic differs significantly, Fig. 10.

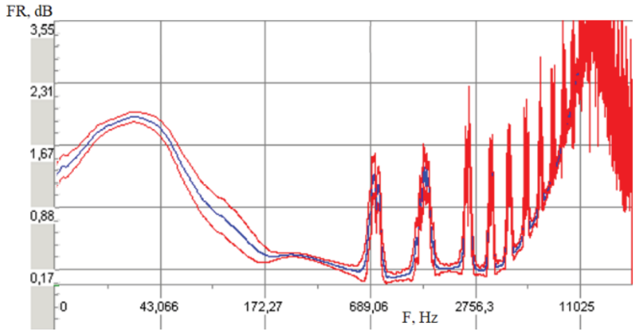

Fig. 10. FR of the equalizer is formed by changes in the instantaneous spectra of the real signal (blue -average value FR, red - dispersion of values FR)

In the primary (before the transmitters) and secondary (to the subscriber) distribution paths, compact signal representation formats, such as MP3, are increasingly being used [6, 9]. In this case, transmitting a test sinusoidal signal requires only a small amount of information. The quality of the SBS will be determined by the transmission rate, and constructing the FR will present certain challenges. During the elimination of psycho-physical redundancy, some spectral components will simply be discarded as they are not perceived by the listener.

Therefore, when forming the FR, an averaged "statistical trend" is established, and deviations from it are assessed during transmission. The smaller the amount of transmitted information, the greater the dispersion of deviations. Fig. 11 provides an example of assessing the FR of the transfer coefficient at an MP3 rate of 64 kbps, formed from a program with a duration of 30 minutes.

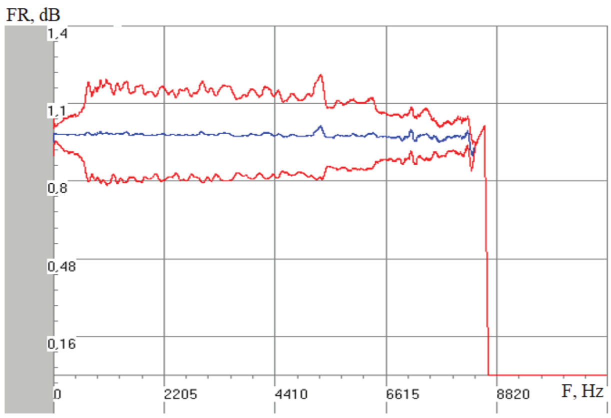

Fig. 11. FR of the path using the MP3 format at a speed of 64 Kbit/s (blue - average value FR, red - dispersion of values FR)

As shown by subjective statistical testing (SST) for assessing the quality of SBS [10], the greater the variance of deviations, the lower the quality assessment of the transmission. Thus, the primary and secondary distribution paths also contribute to the distortions of the FR

Currently, the secondary distribution path does not include subscriber acoustic systems, even though they largely determine the listener's assessment of the quality of the transmitted audio signal.

It is generally accepted to evaluate the quality of acoustic systems (AS) using a harmonic signal that varies in frequency. Fig. 12 shows the FR of the KRK Rokit Powered 5 active near-field studio monitor. It is considered that the provided non-uniformity is ±3 dB in the range of 70 Hz to 20 kHz, which is quite a high standard.

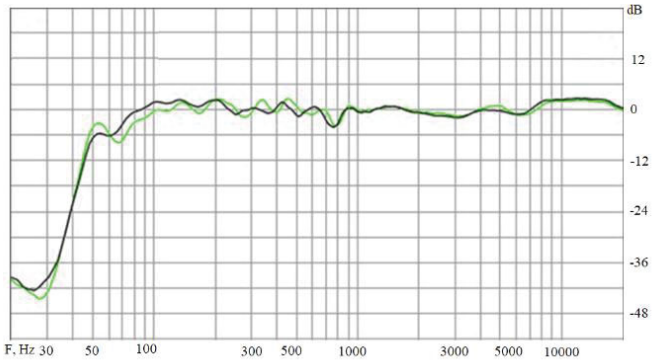

Fig. 12. FR of the active near-field studio monitor KRK Rokit Powered 5 (black - in the near field, green - in the far field)

Acoustic emitters are mechanical oscillatory systems that respond differently to single sinusoidal and multi-frequency excitation [11]. The oscillatory system of the monitor cannot operate equally effectively on a single sinusoidal signal and a multi-frequency signal. An assessment of the FR was conducted for a white noise signal. The results of this assessment are presented in Fig. 13.

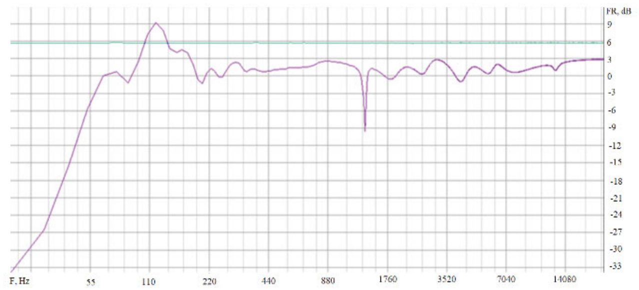

Fig. 13. FR of the KRK Rokit Powered 5 near-field monitor with white noise

As we can see, the unevenness has increased significantly and reaches 10-12 dB. Of interest is the unevenness of the FR when the speakers emit a real information signal. Fig. 14 shows the FR formed by changes in the instantaneous spectra of the signal.

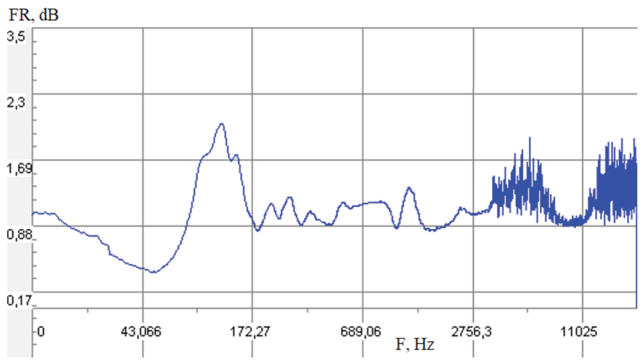

Fig. 14. FR when passing a real information signal

There is a difference in non-uniformity when comparing the emission of a white noise signal to that of a single sinusoidal signal, with the former being lower but higher than the latter. It can be assumed that the FR will change depending on the spectral composition of the signal and its degree of harmonicity, which in turn will alter the listener's assessment of sound quality.

It is known that the FR of the same romance sung with different emotional states—anger, joy, horror—differs significantly. The various deformations of the FR, determined by the nature of the piece, will lead to recommendations for listeners to use specific acoustic systems for reproducing works of certain genres.

III. Conclusion

Monitoring the spectrum of the transmitted signal is the most common measurement operation performed on the SBS. It is carried out during the quality control of transmission, the effects of sound engineering and audio processing, and the assessment of the performance of sound-emitting systems. To date, virtually all paths included in the audio transmission channel are adaptive in both analog and digital implementations. During the adaptation process, the properties of the SBS are altered either in accordance with the capabilities of the transmission channel, to reduce the volume of transmitted information, or to change its emotional informativeness. In all cases, this processing is applied to the real SBS signal, and it is not necessarily recorded during the transmission of test signals. This raises the task of precise spectral analysis of the audio broadcasting signal and the formation of the actual transfer coefficient based on changes in the instantaneous spectra of the SBS.

It has been shown that when assessing the FR of the audio broadcasting channel, significant differences occur when testing it with stationary sinusoidal signals compared to the transmission of real multi-component audio signals, where the distortions arising during the transmission of such real signals remain undetected, affecting the accuracy of the quality assessment of the transmission channel.

It has been noted that the main changes in the transfer coefficient FR in the program formation path are determined by audio processors. An example of FR distortions arising from the use of the Optimod audio processor from Orban is provided, which do not manifest on test signals but are well identified on real broadcast signals, with the cause of these distortions related to multi-band signal processing. It is also shown that such deficiencies are absent in multi-band AGC systems.

Attention is drawn to the distortions that occur during the correction or deformation of the frequency characteristic of the audio signal, where the frequency characteristic of the equalizer, controlled by a sinusoidal signal, confirms its compliance with the interface; however, when checking the operation of this equalizer on a real audio signal, this characteristic turns out to be significantly different.

It is shown that in the primary and secondary distribution paths, the main cause of FR distortions is sections with compact signal representation, such as when using MP3, where an averaged "statistical trend" is formed when constructing the FR, along with deviations from it during transmission, and the resulting distortions only manifest on real signals.

It is proposed to include endpoint subscriber acoustic systems in the secondary distribution path, as they largely determine the deterioration in the listener's assessment of audio signal transmission quality, even when other elements of the path have high-quality indicators.

It is noted that acoustic emitters, which are mechanical oscillatory systems, respond differently to single test sinusoidal and multi-frequency excitation of the real audio signal. Therefore, depending on the spectral composition of the signal and its degree of harmonicas, the FR will change, affecting the listener's assessment of sound quality. Moreover, the specifications for acoustic systems formed during the transmission of a single frequency signal often prove inadequate for the real characteristics of these systems when transmitting a multi-frequency signal.

A method for spectral analysis of audio signals is proposed, which allows for the accuracy and resolution to approach previously unattainable capabilities of the human auditory analyzer, thereby enabling the formation and control of the FR directly during the transmission of the audio program for all devices within the channel.

A method for forming accurate assessments of instantaneous spectra of audio signals over short time intervals is proposed, based on the use of the discrete cosine transform (DCT) with a low level of side lobes and the previously unavailable ability to represent the signal in a complex form.

According to the research results, in all cases, the proposed method for forming the FR assessment has confirmed its effectiveness.

REFERENCES

[1] V. A. Abramov, O. B. Popov, V.G., T. V. Chernysheva, A.A.Borisov “Method and Device for High-Precision Measurement of the Spectrum of Information Acoustic Signals”, RU Patent 2808156, Nov. 24, 2023.

[2] 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: Apr. 19, 2022.

[3] V. A. Abramov, O. B. Popov, V .G., T. V. Chernysheva, A. A. Borisov “Method and Device for Measuring Spectrum and Cepstral Parameters of Information Acoustic Signals of Television and Radio Broadcasting”, RU Patent 2813684, Feb. 2, 2024.

[4] V. A. Abramov, O. B. Popov, I. V. Vlasyuk, A. V. Balobanov, “Method and Device for Measuring the Spectrum of Informational Acoustic Signals with Distortion Compensation”, RU Patent 2756 934, Oct. 7, 2021.

[5] Sh. H. Bazarova “ Pre-processing of audio signals”, *World Science*, 2020. No. 1 (46), pp 6 -9.

[6] Yu. A. Kovalgin, *Digital radio broadcasting: systems and technologies*. Hotline - Telecom. 2021, 580 p.

[7] P. Záviška, P. Rajmic, A. Ozerov and L. Rencker, "A Survey and an Extensive Evaluation of Popular Audio Declipping Methods," in IEEE Journal of Selected Topics in Signal Processing, vol. 15, no. 1, pp. 5-24, Jan. 2021, doi: 10.1109/JSTSP.2020.3042071.

[8] S. J. Schlecht, L. Fierro, V. Välimäki and J. Backman, "Audio Peak Reduction Using a Synced allpass Filter," ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Singapore, Singapore, 2022, pp. 1006-1010, doi: 10.1109/ICASSP43922.2022.9747877.

[9] B. Ming and P. Wu, "Research on Audio Signal De-noising and Simulation Processing," 2019 International Conference on Communications, Information System and Computer Engineering (CISCE), Haikou, China, 2019, pp. 192-194, doi: 10.1109/CISCE.2019.00051.

[10] Recommendation ITU-R BS. I 284-2 (01/2019). General methods for the subjective assessment of sound quality. BS Series. Broadcasting service (sound).

[11] T. Bai, L. Xie, Z. Li, J. Yang, Z. Chen and P. Wan, "A High-Precision Audio Z-Δ D/A Converter," 2020 IEEE 14th International Conference on Anti-counterfeiting, Security, and Identification (ASID), Xiamen, China, 2020, pp. 120-123, doi: 10.1109/ASID50160.2020.9271769.