Changing the Properties of the Audio Broadcast Signal in Adaptive Transmission Channels

Oleg Popov, Tatiana Chernysheva, Andrey Borisov,
Pavel Sapronov, Kirill Orlov
MTUCI
Moscow, Russia
olegp45@yandex.ru, krba2012@yandex.ru, a.borisov199@gmail.com,
Pavelozone99@gmail.com, kvororloff@yandex.ru

Abstract—A paradoxical situation has occurred and is maintained, when all existing channels of sound broadcasting signal (SBS) are adaptive and are determined by the capabilities of matching the properties of the transmitted signal with the capabilities of the transmission channel, and the quality assessment methods are designed for a stationary channel with unchanged properties in the process of transmission. We measure the change of the signal form, the simplest stationary test signal, and for the real signal the form is not known to be preserved. The following are examples of the links that make the largest changes in the SBS.

I. Introduction

There are no audio transmission channels left today that can keep the shape of the real signal being transmitted. In analog transmission systems, such waveform change is determined by automatic level controllers (ALC), audio processors, compander systems, noise suppressors, equalizers and a whole range of devices that help make the signal "preferable" to the listener. When checking the quality of transmission through such channel, the creators of the recommendations and State Standards carefully recommend to disconnect all the "devices for additional processing" and to check the

preservation of the form of the stationary harmonic signal, and then by connecting them to a channel with completely different properties to transmit the real signal. By now it is not difficult to form an error signal by the real signal at the channel input and output and to remember that permissible

distortions introduced in GOST were formed exactly by the real signal [1]. Even more changes a form of a signal in modern digital channels. In addition to the distortions determined by the method of signal representation, there are distortions due to the change of sampling frequency in the transmission channel, audio processing [2], and elimination of statistical and psychophysical redundancy. The main sources of SPL distortions in the transmission channel are considered below.

II. Changes in the audio broadcast signal on alternate channels

Changes in SBS are not always associated with signal spectrum enrichments, so it is not quite correct to call them distortions. As already mentioned, a number of static devices are used in the transmission channel, such as pre-distortion and restoration loops, compressors and dynamic range expanders (compressors and expanders), as well as a variety of ALC, which recently transformed into the category of "audio processors".

The patterns of signal changes in all of these devices are about the same, so let's look at the example of signal changes in an automatic level control. The signal is subjected to automatic regulation already in the programming path, almost immediately after the microphone. Already in the studio it is regulated several times, and then at least at the input of the primary and secondary distribution path. At each stage the regulation is done, according to the control signal that is generated from the sound signal itself, and therefore even theoretically it is impossible to restore the form of the original signal. Fig. 1 shows the oscillograms of the audio broadcast signal at the input and output of the level autoregulator. On the abscissa axis is time, and on the ordinate axis - level in steps of quantization.

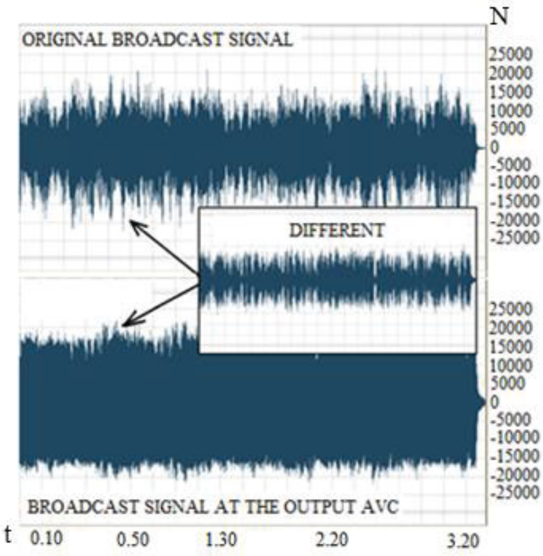


Fig 1. Oscillograms of the signal at the input and output of the autoregulator, as well as the difference between them (error signal).

The scale is the same for the audio broadcast signals and the error signal, so you can see that the signal variation is much higher than the 2% allowed by State Standard. Fig. 2 shows the spectrum of the difference signal, from which we can see that the error signal spectrum is practically not masked by the signal spectrum, which does not allow to reduce the visibility of noise and interference. As a result of companding, it is possible to reduce the level of noise in the pause of the audio broadcast signal, but the signal-to-noise ratio is practically unchanged by the auto-adjustment system in the background of the signal. The companding system allows you to increase the signal-to-noise ratio in the pause by about 20 dB according to subjective perception [3].

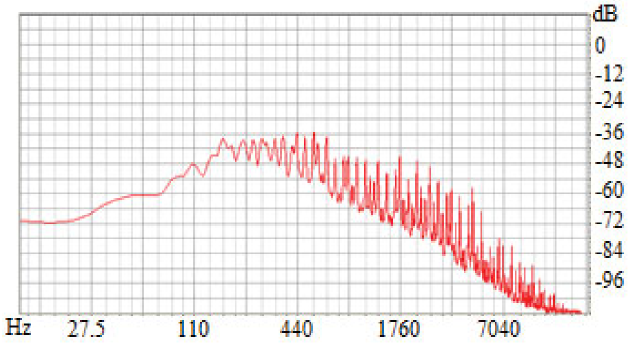


Fig. 2. Error signal spectrum at the AVC output.

Compounder systems are very sensitive to changes in the transmission factor of the path, which is not covered by State Standard, but in many ways it determines the quality of the transmission. To adequately regulate the SBS at the input of the compander system, it is desirable to install ALC, which maintains the original diagram of signal levels.

II. AUDIO BROADCAST SIGNAL DISTORTIONS IN DIGITAL TRANSMISSION PATHS.

As is known, 80% of the signals belonging to the musical culture of mankind are low-level signals. Distortions of these signals are particularly noticeable to the listener because at signal levels below 40 dB the effects of frequency and time masking stop working, allowing to hide noise and distortion under the threshold of audibility of the auditory analyzer excited by the main signal.

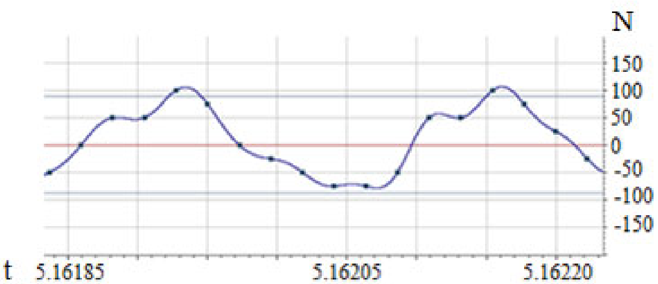


Fig. 3. Normalized raw low-level signal

An example of an oscillogram of such a signal recorded at -70 dB with a 16-bit representation is shown in Fig. 3. The visually and hearing analyzer are the same, so if we visually see that the sine is slightly non-sine, we will also notice the distortion audibly. In this case, their value is about 15%, with a threshold visibility for the average listener of 2-3%.

The amplitude-frequency response of the given signal is shown in Fig. 4.

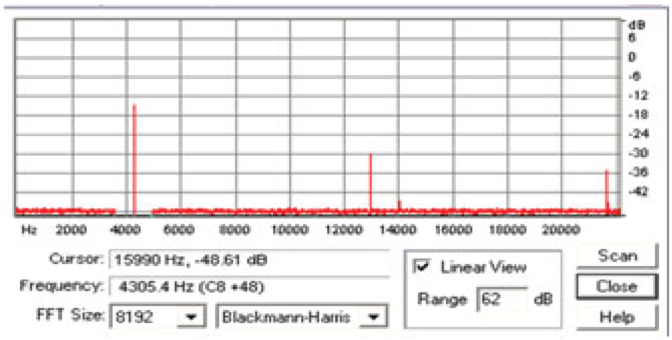


Fig. 4. Spectral analysis of the original low-level signal

Such distortions are characteristic of analog-to-digital converters with bit-wise weighting. When a high intermediate frequency sigma-delta modulator is used in the analog-to-digital conversion, these distortions can be avoided, but only for a stationary harmonic signal. Fig. 5 shows an oscillogram of a signal at the same -70 dB level, synthesized in a modern sound editor using a sigma delta AlC.

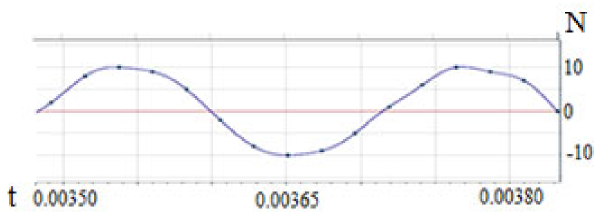


Fig. 5. A low-level signal digitized with a sigma delta ALC.

The absence of distortions is determined by the effective work of the interpolators in the digital-to-analog converter, which increase the sampling frequency, the so-called "oversampling". Unfortunately, the interpolation used in the time domain is effective only for stationary signals with a spectrum, which are used as measuring signals when evaluating sampling rate changing devices. When digitizing a real signal with quasi-stationary spectra on the durations of individual audio objects the distortions reappear, which is especially noticeable when the sampling rate in the transmission channel is changed.

Distortions occur not only during the analog-to-digital conversion, but also during transmission through the channel as a result of the need to change the sampling rate. Representation errors are well masked by the signal itself, although on a subconscious level they are negatively perceived by the listener. Fig. 6 shows oscillograms of the signal with the most common variant of sampling rate change 48-32-48 kHz.

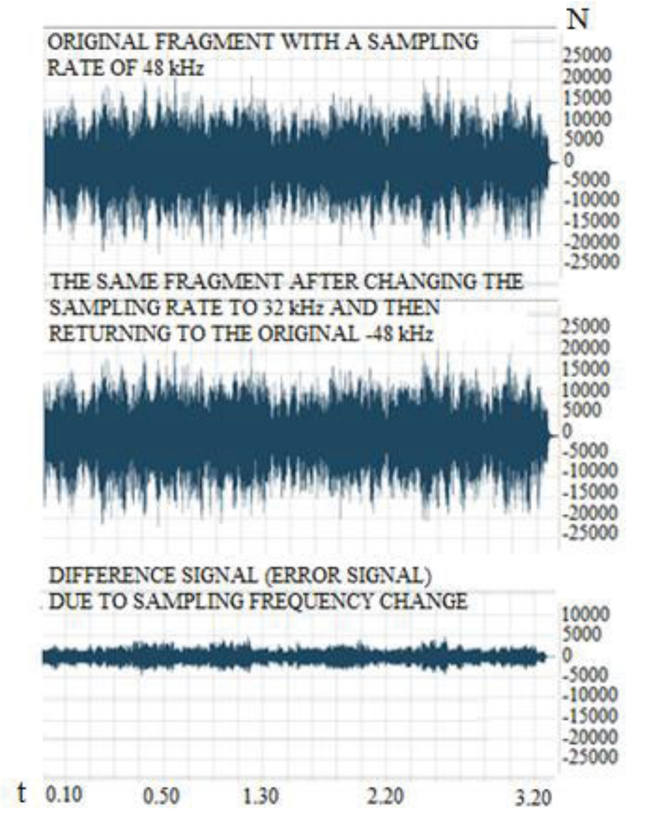


Fig. 6. Oscillograms of SBS with a sampling frequency change of 48-32-48 kHz.

Errors arising in the process of changing the sampling frequency have interested channel operators relatively recently due to the increased demands of listeners to the quality of the transmission. As an example, Fig. 3 are the oscillograms of SPL, in which the sampling frequency first decreases from 48 kHz (studio) up to 32 kHz (the primary distribution channel), and then increases again up to 48 kHz in the other radio house. The situation is quite typical for the distribution of radio programs across the country. The amplitude-frequency characteristic of the error signal is shown in Fig. 7.

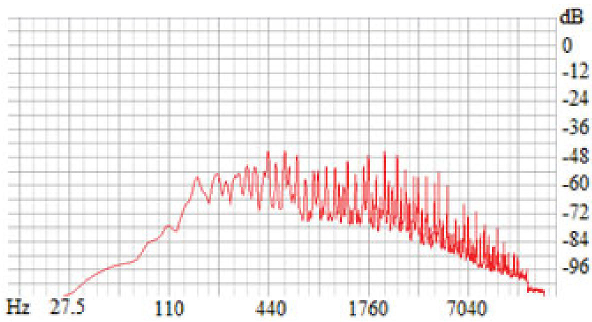


Fig. 7. Spectrum of the error signal when the sampling rate changes

The high level of spectral components of the error in the area of high frequencies, which knowingly will not be masked by the signal spectrum, can be clearly seen.

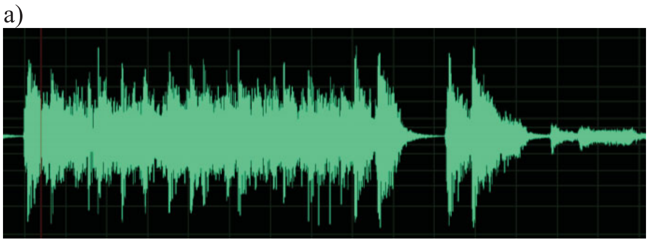
Analyzing the oscillogram of the oversampling error signal, you can see that the error increases at moments of non-stationarity of the signal, leading to changes in its spectrum (attack). That is, the signal is distorted at moments when these distortions are particularly noticeable to the listener. Recall, that the change of signal rise process is noticed by the listener at 0.3 ms. duration [4]. The presence of these distortions is determined by the way of formation of discrete samples of a new sequence in the points where they were not originally. This example uses the sampling rate change algorithm integrated in the sound editor Adobe Audition, which works according to the traditional scheme: sample rate increase, interpolation of intermediate samples, decimation (discarding) of unnecessary samples. Such an algorithm works well on a stationary harmonic signal, which is used to evaluate the channel quality. Interpolation errors arise when sound objects change and are especially noticeable at the moment of quick appearance of a new sound object.

It is possible to avoid such distortions by using the sampling frequency conversion algorithm developed at the Department of Television and Sound Broadcasting (TaSB) of MTUCI [4]. In the algorithm the interpolation is performed in the frequency domain and especially in the frequency domain, where the signal energy of the sound broadcasting is known to be absent. The error signal level does not exceed -92 dB and, practically, is determined by the accuracy of the calculations.

Unfortunately, in SPL transmission and storage channels, traditional sampling rate changers with interpolation of intermediate values in the time domain are most often used. Our calculations have shown that even for a stationary sinusoidal signal it is necessary to use an interpolator based on a nonrecursive digital filter, including about 1000 delay lines and multipliers to meet GOST requirements for allowable distortions. And the interpolation error will sharply increase at the junction of different audio objects with different quasi-stationary spectra.

In the case of using sigma delta modulation it is possible to reduce the distortions at the quasi-stationary parts of the signal, but they still remain at the moments of non-stationarity. Fig. 8 a, b, with preservation of scale, shows the original SBS and the error signal after transformation of sampling frequency 48-32-48 kHz.

a)



б)

Fig. 8. Oscillogram of the original signal and the error signal after changing the sampling rate.

Fig. 9 a, b, in an enlarged scale shows an oscillogram of the original signal and the error signal enlarged in scale, you can see the synchronization of errors with the moments of changes in the amplitude of the signal.

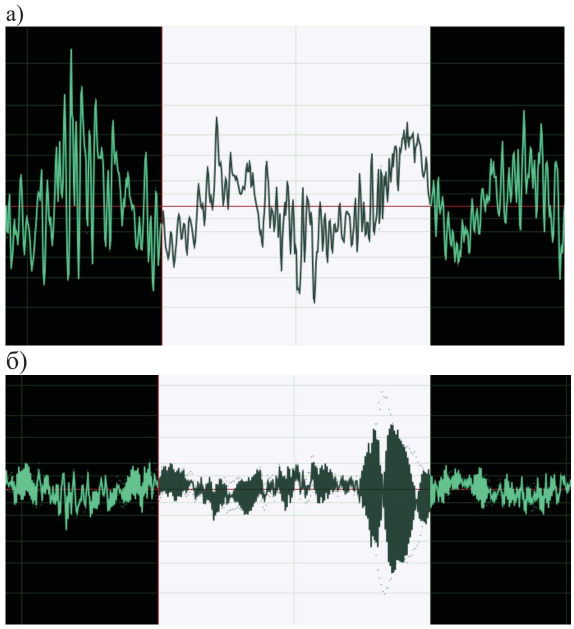


Fig. 9. Oscillograms of the original signal and enlarged error signal.

Therefore, using sigma delta modulation instead of the traditional PCM representation enables a better compliance with the standards for intolerable distortions during measurements in accordance with the existing methods of quality evaluation on stationary low-component signals, but does not provide a significant gain on the real signal.

III. DISTORTIONS OF THE AUDIO SIGNAL IN REDUNDANCY-ELIMINATED TRACTS.

When transmitting and storing SBS, they seek to reduce the volume of the signal using its statistical and psychophysical redundancy.

Psychophysical redundancy, determined by the processes of signal perception by the listener, is used in the method of instantaneous compounding (IC). IC encodes the signal at different quantization steps depending on the level, with a small step in the area of low levels and a large one in the area of high levels. The resulting noise is proportional to the quantization step and is successfully masked by the signal itself through frequency and time masking in the auditory analyzer, i.e., psychophysical redundancy is eliminated.

In the course of development, a method of almost instantaneous companding (AIC) was developed, in which redundant statistical redundancy is preserved, redundancy in signal description. Knowing that the volume (level) of the signal changes relatively slowly the step size of the quantization can be transferred not with each reference as in IC, but once every one millisecond.

Quasi-logarithmic compression characteristics of type A (in audio broadcasting) and type μ (in telephony) are widely used.

The selection of the quantization step is based on the maximum signal, the spectral component, with all other spectral components modulated by this variable quantization step.

Fig. 10 shows the error signal obtained by subtracting the compounded output signal from the input signal. The signal is generated for the speech signal.

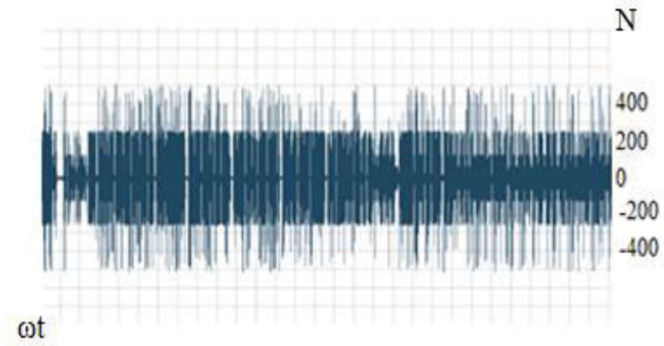


Fig. 10. Companding error signal.

We can see that the error signal has a discrete character, conducting its spectral analysis, we can see that it is almost uniform in the entire frequency spectrum. As a result, for the vocalized vowel sounds, with the predominance of low frequencies.

As a result, during the sounding of vocalized vowel sounds, mostly low-frequency, high-frequency noise is not masked by the signal due to frequency and temporal masking.

The spectra of the original and compounded harmonic signal are shown in Fig. 11a, 11b. It can be seen that even a single stationary harmonic signal is enriched in the MC process due to amplitude modulation by varying the quantization step determined by the signal amplitude.

Not convenient for coding with IC is a biharmonic signal, shown in Fig. 12. Due to the fact that the quantization step is determined by the amplitude of mainly low-level signal, at maximum amplitudes the quantization step is too large to describe a small signal and it disappears for some time. That is, the high-frequency low-level sound acquires the amplitude modulation shown in Fig. 13a,b for the signal: $X(t) = a\mu \sin 2\mu f_l(t) + b\mu \sin 2\mu f_u(t)$, where $f_l << f_u$ and $a >> b$.

We can see that the low-level signal has acquired lateral spectral components determined by the frequency of the powerful low-frequency component.

Similar distortions will be observed at almost instantaneous companding of AIC. When a real multicomponent signal is transmitted, each spectral component will acquire side frequencies determined by the maximum spectral component in amplitude.

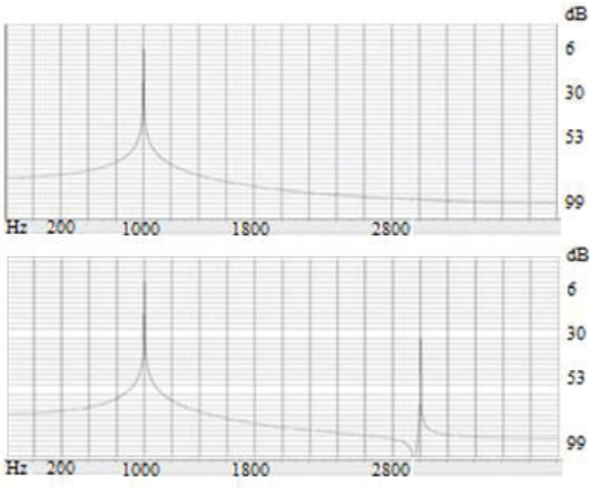


Fig. 11. Spectra of the original and composite signals

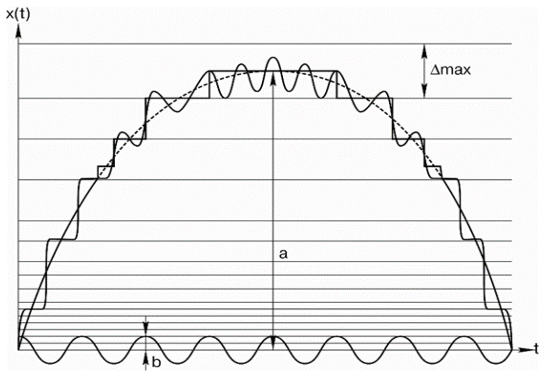


Fig. 12. Quantization of a bi-harmonic signal

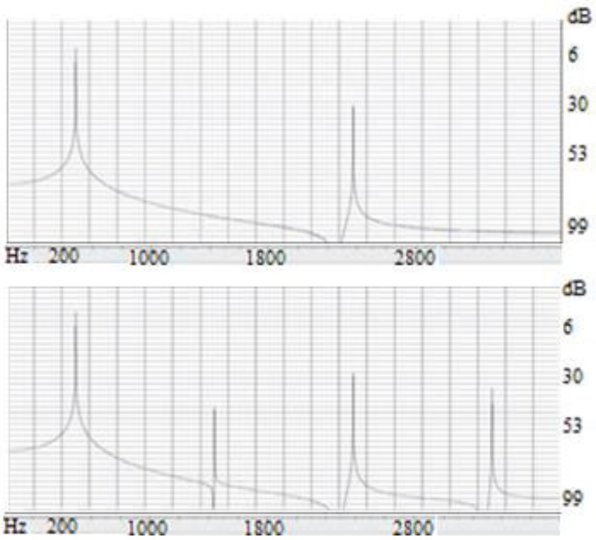


Fig. 13: Distortion spectrogram of a bi-harmonic signal

AIC is used for coding subband signals in MP3 format like the difference coding methods (DPCM). Distortions for multicomponent signals at such coding are in many respects similar, that is why in Fig. 15, 16 the spectrograms of the signal after the difference coding of a five-component signal the spectrogram of which is given in Fig. 14. Four bits were used to describe the difference.

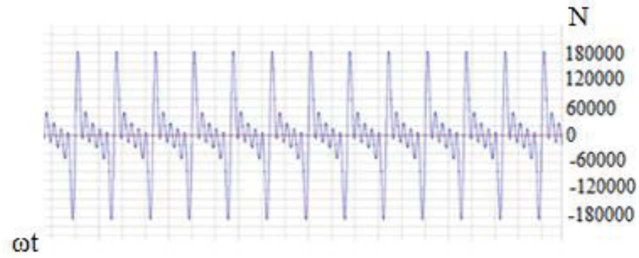


Fig. 14. Test signal of five harmonics

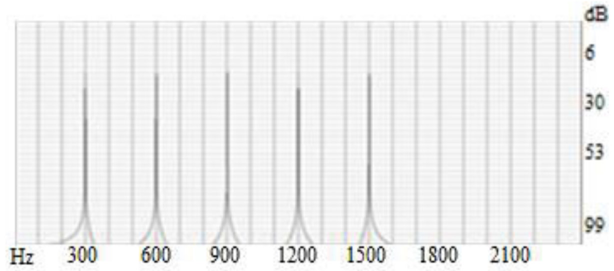


Fig. 15. Spectrum of the initial signal of the five harmonics

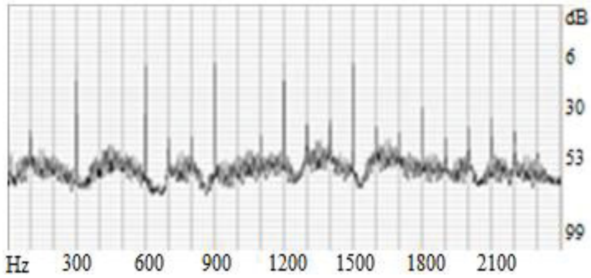


Fig. 16. The spectrum of the test signal in the difference representation

As we see for IC, AIC and DPCM the signal spectrum is enriched with a spectrum of modulation noises. It is believed that these noises are invisible to the listener due to frequency and time masking in the auditory analyzer. It should be noted that these effects have been studied for the perception of narrowband noise signals and low-component harmonic components. It is noted [1] that the regularities of perception of a real signal significantly differ, including a different steepness of the decline of frequency masking in the low and high frequencies, differences in the threshold perceptibility of the correlated audio signal, which is actually noted in the normalization of the perceptibility of the audible transient interference. In the process of research of listeners' estimation of signals of compactly presented SBS a confident preference of sound with the minimum spread of spectral components transmission coefficient is noted. The reduction of the digital stream rate required to transmit the SBS under conditions of limited path capacity first affected the secondary distribution paths, digital radio broadcasting, Internet broadcasting, then the expensive primary distribution channels, and then the program formation path (studio), as well as the widely used systems for storing archives. The compact representation of SBS does not intend to preserve the form of the signal, and provides only high subjective quality of transmission. The difference between the original and compactly presented signal is big enough and can be up to 15-20% of the original signal (Fig. 17). At the same time in the evaluation of a stationary harmonic signal, according to GOST P52742-2007 Sound broadcasting channels and sections.Typical structures. Basic quality parameters. Methods for measuring distortion is absent. Fig. 17, 18 shows the oscillograms of the SBS before and after encoding and the error signal.

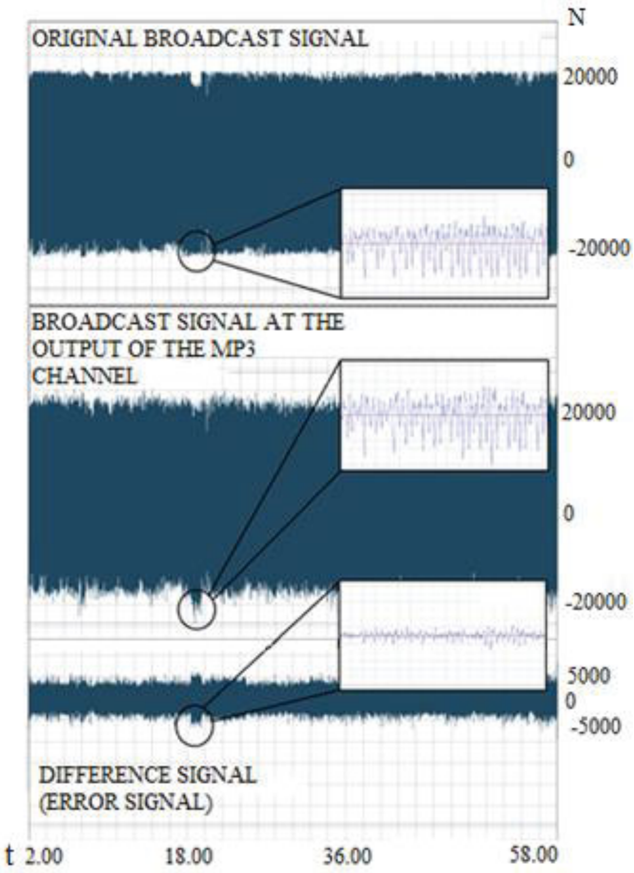


Fig. 17. The SBS at the input and output of the MP3 encoder, as well as the error signal.

The spectrum of the error signal is shown in Fig. 16.

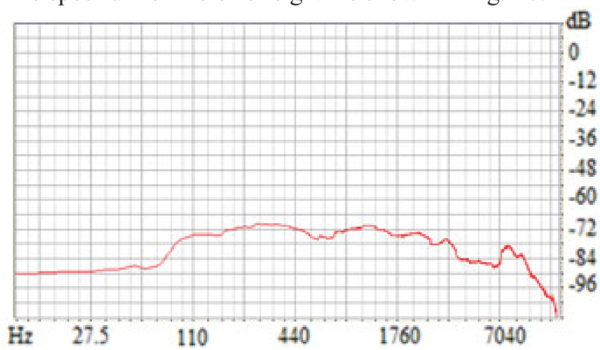


Fig. 18. The spectrum of the error signal in the compact representation of the SBS

Analyzing the oscillograms of the error signal, we can notice that the error value is variable along the spectrum and is especially great at the areas of signal non-stationarity - attacks. It is also modestly mentioned in the international recommendations for the IC format on 30 pages of 32 pages of the document: "There are also some distortions in the signal rise and fall. Recall that the rising elements of the signal (attacks) are related to the intelligibility of the recognition of the musical signal by the way of sound extraction, and the decline reflects the conditions of the recording, that is, the process of reverberation in the room. The adaptive capacity of the human ear is very great and, for the most part, our listeners have adapted to sound with such distortions, sometimes being surprised by the difference between compactly presented and natural sounds. Most methods of compact representation use signal separation into sub-bands and subsequent coding. In this case, after mixing the signal back, the rise and fall process is no longer determined by the properties of the signal, but by the properties of the filter. This effect is especially pronounced when transmitting over a channel that includes several paths with a compact representation of the SBS.

Fig. 7 shows oscillograms of harmonic steady-state signal when transmitting over a channel with a compact representation at 32 kbits/sec, when exactly by the change of such a signal the quality of transmission over the channel is evaluated.

We can see that the channel provides an excellent transmission quality and can be used for SRS transmission, when measured according to the non-designated GOST 11515 Sound broadcasting channels. Basic quality parameters. Methods for measuring, according to the method transferred to GOST P52742-2007 Sound broadcasting channels and sections.Typical structures. Basic quality parameters. Methods for measuring.

The spectrum of the error signal shows the presence of high levels of this error in the high frequencies. Therefore, it will be badly masked by the signal itself and thus will be noticeable to the listener. This is confirmed by the results of subjective statistical tests done at the MTUCI TaSB Department on investigating the perceptibility of changes in SBS during the transition from the representation in PCM format to IC format with different transmission speeds. According to the results of examinations carried out according to international recommendations with involvement of up to 25-years-old MTUCI students as experts, the threshold of SBS changes visibility is about 256 kbps per mono channel, what pleased MTUCI teachers a lot.

IV. Conclusion

To date, distortions, or rather changes in the signal in the channel of audio broadcasting signal are very large and are not covered by existing techniques for measuring objective parameters of the signal.

As shown in this article, changes or distortions occur already at the stage of analog-to-digital conversion and continue at all stages of audio broadcasting signal transmission through the transmission or storage channel and, unfortunately, are not always taken into account by the designers of transmitting systems.

Numerous devices and algorithms to improve the sound quality, previously used only in the program formation path: enhancers that increase the steepness of attacks, maximizers that increase the relative average power and presence effect, vitalizers that animate the signal, exciters that increase the prominence of short signals, are now distributed across all paths of the transmission channel, providing an increase in the quality of the SBS transmission for the listener. Separate mention should be made of the devices for "finishing" the signal directly before the transmitting device. Such devices are used practically by all broadcasters, interested in the number of listeners. Devices in many ways compensate for SPL losses in primary distribution channels due to compact design, changes in sampling frequency, analog and digital compandor's. To some extent these devices imitate director's signal processing in the channel of program formation, therefore it is desirable that radio station sound engineer should take part in their tuning, but unfortunately it is a rarity. Especially special are the developments of the TaSB MTUCI department which provide undistorted companding in channels with insufficient dynamic range [6,7,8].

Variations of SBS in the transmission channel are large and are not covered by existing methods of objective rationing. Moreover, the methods purposefully introduce the conditions of quality parameters measurement, not corresponding to the operating conditions, for example "before conducting measurements, turn off all devices of additional signal processing", which means switching off all audio processors, noise suppressors, enhancers, vitalizers, etc., and after their connection the signal is transmitted through the channel with completely different properties. As a rule, test signals are stationary and low-component ones that are not distorted in the transmission systems with the redundancy elimination as opposed to the quasi-stationary multicomponent real signal.

The only method to evaluate the quality remain subjective statistical tests that are not applicable when there is a need to quickly assess the quality of transmission and its changes, for example, at the expense of new ways to coding or audio-processing. There is an urgent need to study and identify the signal parameters, changes of which determine the listener's evaluation of the transmission quality. It is necessary to involve sound engineers, musicians in such work, ensuring the possibility of predicting the listener's evaluation of the sound quality by the results of changes in the objective parameters of the SBS.

References

[1] O.B. Popov., V.A. Abramov., S.A. Litvin, T.V. Chernysheva «Increasing the resolution capacity of spectral analysis on short time intervals» Proceedings of the XV International Branch Scientific and Technical Confer-ence "Information Society Technologies". (3-4 March 2021 Moscow, MTUCI). M.: MTUSI, 2021 P.90-92.

[2] V. A. Abramov, O. B. Popov, T. V. Chernysheva, V. G. Taktakishvili and A. A. Ovchinnikov, «On-board Transmission Quality Assessment Using Short Audio Signal» 2021 Systems of Signals Generating and Processing in the Field of on Board Communications, 2021, pp. 1-6, doi: 10.1109/IEEECONF51389.2021.9416085.

[3] V.A. Abramov, V.G. Taktakishvili, A.A. Ovchinnikov, O.B. Popov «Objective Assessment of Transmission and Informativeness of a Speech Signal According to Statistical Parameters» Proceedings of the FRUCT 24, Moscow, Russia, 8-12 April 2019, p. 741 – 746.

[4] V. A. Abramov, O.B.Popov, S.G. Richter. «Audio processing of digital broadcasting signals and its consequences» T-Comm. Telecommunications and Transport 2016, No. 6, p.17-20.

[5] V.A. Abramov, O.B. Popov, V.G. Orlov, "Method and apparatus for companding audio broadcast signals", RU Patent 2 691 122, Jun. 11, 2019.

[6] 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.

[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, V.G. Taktakishvili, «Method and apparatus for companding with predistortion of audio broadcast signals», RU Patent 2 731 602, Apr. 9, 2020.