

Measuring Distortion in Amplifiers

Introduction (instead of summarizing)

It is known that in the products of distortions of a sinusoidal signal in amplifiers, predominantly harmonic components appear that are multiples of the fundamental frequency f , where $n = 1, 2, 3, \dots \infty$. These distortion products can be isolated in steady state using spectrum analyzers, as well as in summary form using THD meters.

By measuring THD with sinusoidal signals, we can only measure what the amplifier has added, but we cannot measure what it has lost. In fact, the lost information is much more important! And to measure it, completely different testing methods are needed.

Any amplifier (except the ideal one) has a limited bandwidth from above, which imposes, in addition to non-linear distortions, additional linear distortions: phase shift, change in signal amplitude.

If the amplifier has a closed input (non-DC amplifier), then in addition to linear and non-linear distortions, transient distortions occur in the distortion products, which are often confused with linear distortions.

In DC amplifiers, it is often necessary to use a servo control system, which, if implemented incorrectly, can also introduce additional transient distortion and non-linear distortion in the low-frequency region.

Phase and gain margins often require the use of an inductor at the output of the amplifier. The presence of inductance can also be the cause of additional distortion introduced by amplifiers when operating on a real reactive load.

In a number of operational amplifiers, in the speed parameters section, such a parameter as tPD (time Propagation Delay) has been added. On the horizontal sections of the group delay (Group Delay) $tPD = GD$. Developers who understand the meaning of this parameter began to indicate its value in the specifications. In high quality amplifiers, this parameter rarely exceeds 100 ns! and has a constant value up to at least 300 kHz.

An RC circuit time constant with this value corresponds to a cutoff frequency of 1.6 MHz. I've seen rave reviews for 2.2 MHz DC amplifiers on numerous occasions.

In an interview [29], the legendary developer of audio products without the use of negative feedback, Charles Hansen, who is a musician and a physicist by training, says the following:

“The real problem is that there are almost no measurable parameters that correlate with perceived sound quality.

The ear/brain is much more sensitive to time-related information than to any other parameter.

In analog circuits, feedback loops take a time-delayed signal from the output and send it back to the input in an attempt to correct an error that has already occurred. This creates a form of temporal distortion that cannot currently be measured, but is clearly audible. The math (and the test equipment) tells us that the correction is fast enough, but our ears are telling us otherwise.”

Indeed, for periodic sinusoidal signals, the summation of phase-shifted signals gives a sinusoidal signal of exactly the same shape, but a different amplitude. Changing the phase shift changes the amplitude of the signal without affecting its shape. But this is so **ONLY** for sinusoidal signals.

For non-stationary signals of a complex shape, such as musical signals, this is not true, for a complex signal (this is the same as for the first period of the burst of a periodic signal), there is no “tail” of the previous period for correcting the first one. We've arrived!

Moreover, negative feedback does not work well with noise-like signals, and yet human hearing has a logarithmic dependence (Weber-Fechner law) and is very sensitive to the microlevel information of an audio signal at the noise level.

Let's try to understand with the help of simple arithmetic whether the time distortions mentioned by Charles Hansen are so elusive.

And now imagine that we have an ideal amplifier (but we doubt it) and to make sure that there are no introduced distortions, we will measure the vector errors. To do this, it is enough to use the Hafler SWDT method [13] and compare the output signals with each other, Fig. A.

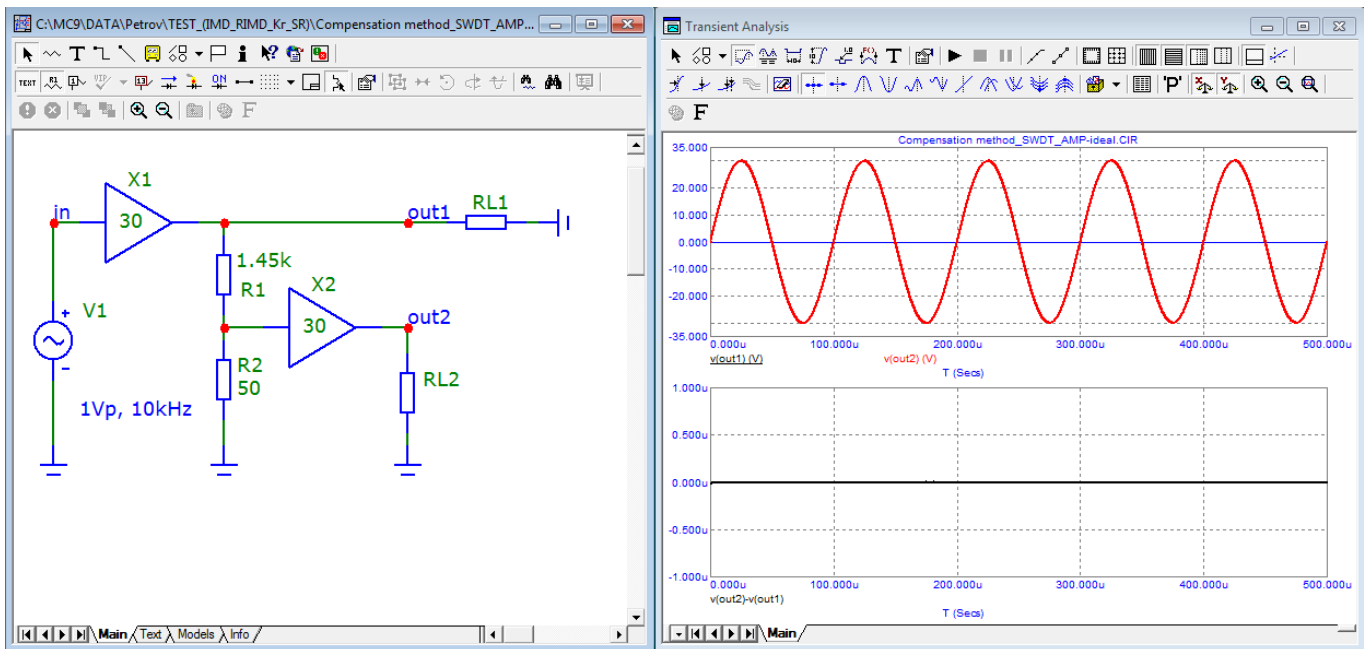


Fig. A. Block Diagram of Hafler Vector Distortion Measurement (SWDT)

As you can see, the difference between the output signals is zero, which corresponds to the absence of any introduced distortions.

And now let's give a signal in the form of bursts, fig. B

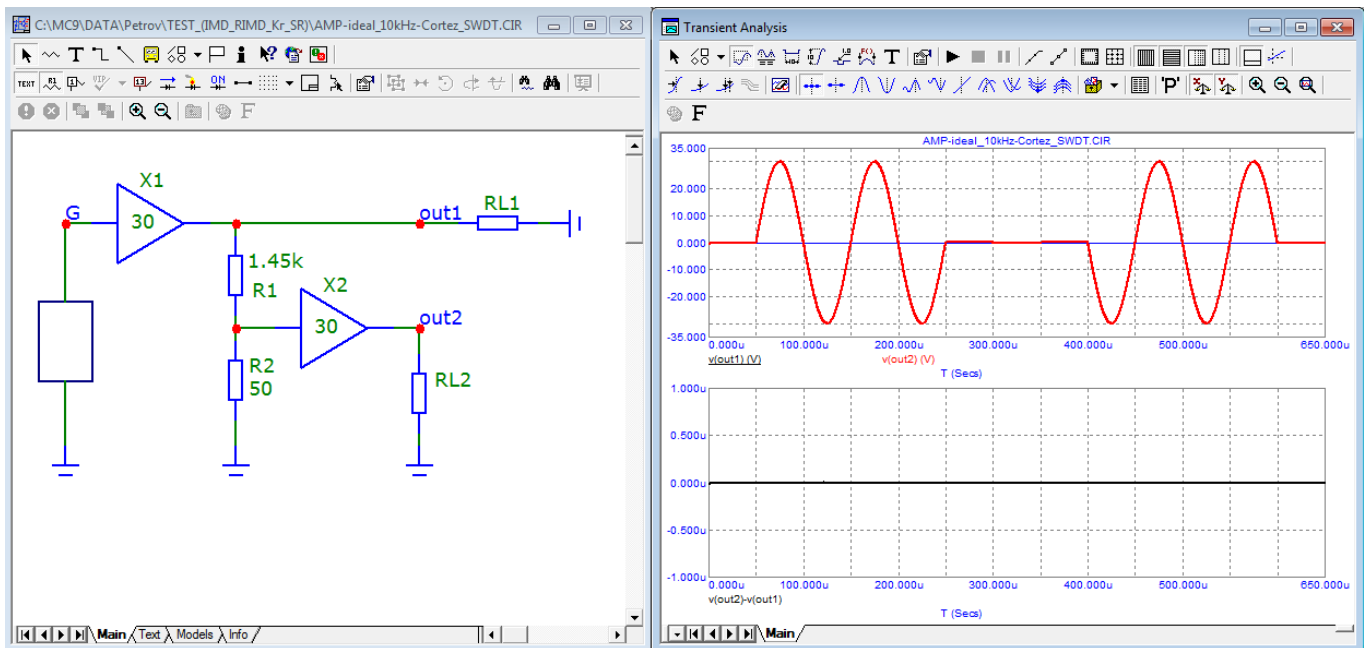


Fig. B. An example of using test signals in the form of bursts

As the test showed, even with such a complex signal at the output of the amplifier under test, there are no additional spectral components that scare in a number of scientific publications. However, remember that we are dealing with an ideal amplifier whose input capacitance is zero, the signal propagation delay (time Propagation Delay) is also zero. This is what is required for the perfect operation of the amplifier, as Cyrill Hammer [20] spoke about.

Now let's move from an ideal amplifier to a block diagram of a real amplifier having a signal propagation delay of 1 μ s and introducing negligible non-linear distortions in the form of the 3rd harmonic with a level of 3 μ V, Fig. C.

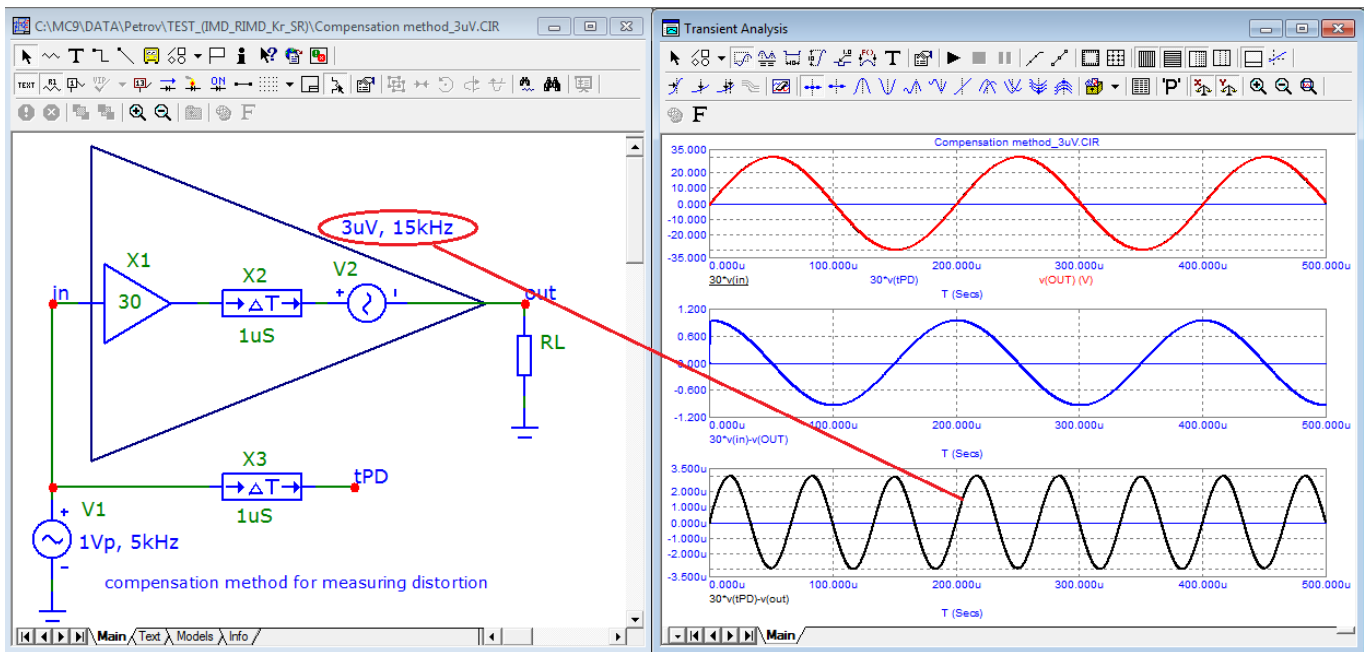


Fig. C. Structural block diagram of an amplifier based on ideal components

Figure C shows a block diagram of an amplifier consisting of three components: an ideal amplifier X1 with $K_u = 30$, an ideal delay line X2 for $1 \mu s$ simulating the delay time of the signal in the amplifier and a generator V2 (15 kHz, 3 μV) simulating the distortion introduced by the amplifier.

Let's apply a signal with a frequency of 5 kHz and a level of 1 V (peak) to the input of the amplifier. In this case, the oscillator signal V2 will be the third harmonic.

We will measure vector errors (SWDT) and all types of distortions of the simulated amplifier using the compensation method.

On the first graph, we will display the output voltage and the normalized input signal and the input signal delayed by the time it passes through the amplifier (i.e., by the time tPD). Signal normalization consists in multiplying them by K_u , i.e. by 30.

Since the delay of the output signal and its "distortion" (3rd harmonic) are negligible compared to the signal period of 200 μs and its amplitude, all three signals merge into one line (top red graph).

The second graph is obtained by subtracting the output voltage from the normalized input voltage and represents the Hafler vector error or SWDT.

For a more accurate measurement of the distortion introduced by the amplifier, it is necessary to subtract the output signal from the normalized tPD signal, which is shown in the third graph. As you can see, the subtraction result gave us the 3rd harmonic in its purest form: essentially the same signal as V2.

Since in the model as an amplifier with $K_u = 30$ an ideal amplifier is used, which has a flat frequency response and phase response and thus does not introduce linear distortions, we did not find them in the distortion products.

Consider a couple of examples on models of real amplifiers in one of which linear distortions also appear. Let us measure the distortion by the compensation method at the output of the BC-1 [30] amplifier model of a signal consisting of sequences of bursts with a frequency of 10 kHz of different amplitudes, fig. D.

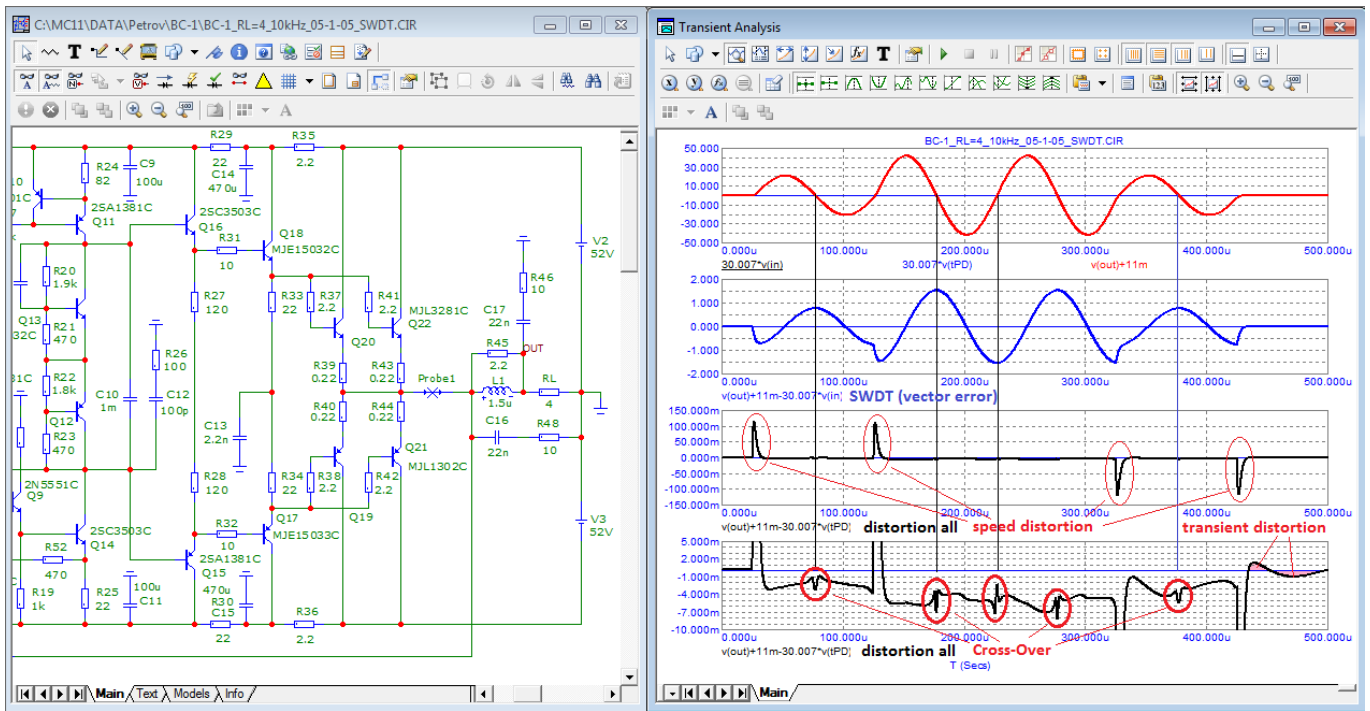


Fig. D. The test result of the BC-1 amplifier model using 10 kHz bursts of different amplitudes (without phase discontinuity).

On the first graph, as in the previous case, three signals are displayed: the output signal and the normalized input and tPD signal.

The second graph (blue) shows a graph of vector errors (SWDT) whose amplitude is 1.5 V, which does not fit into the norm (-60 dB) according to Hafler [13].

The third and fourth graphs show the result of distortion measurement by the compensation method. In the third graph, we see high-speed distortions that occur every time a sudden appearance (disappearance) of voltage, as well as a change in signal amplitude. The amplitude and duration of bursts is nothing more than a loss of useful information! Moreover, relative losses can be much larger for small signals, for example, for higher harmonics of musical instruments that can simply be “blurred” or shifted in phase to another region as a result of amplitude-phase conversion.

In the fourth graph, the distortion products are shown on an expanded scale. Since the amplifier has a relatively large signal propagation delay ($t_{PD} = 215$ ns), the feedback does not cope well with switching distortion.

Let's carry out a similar test of the amplifier model [31], in which the signal propagation delay time is an order of magnitude smaller ($t_{PD} = 24$ ns), fig. E.

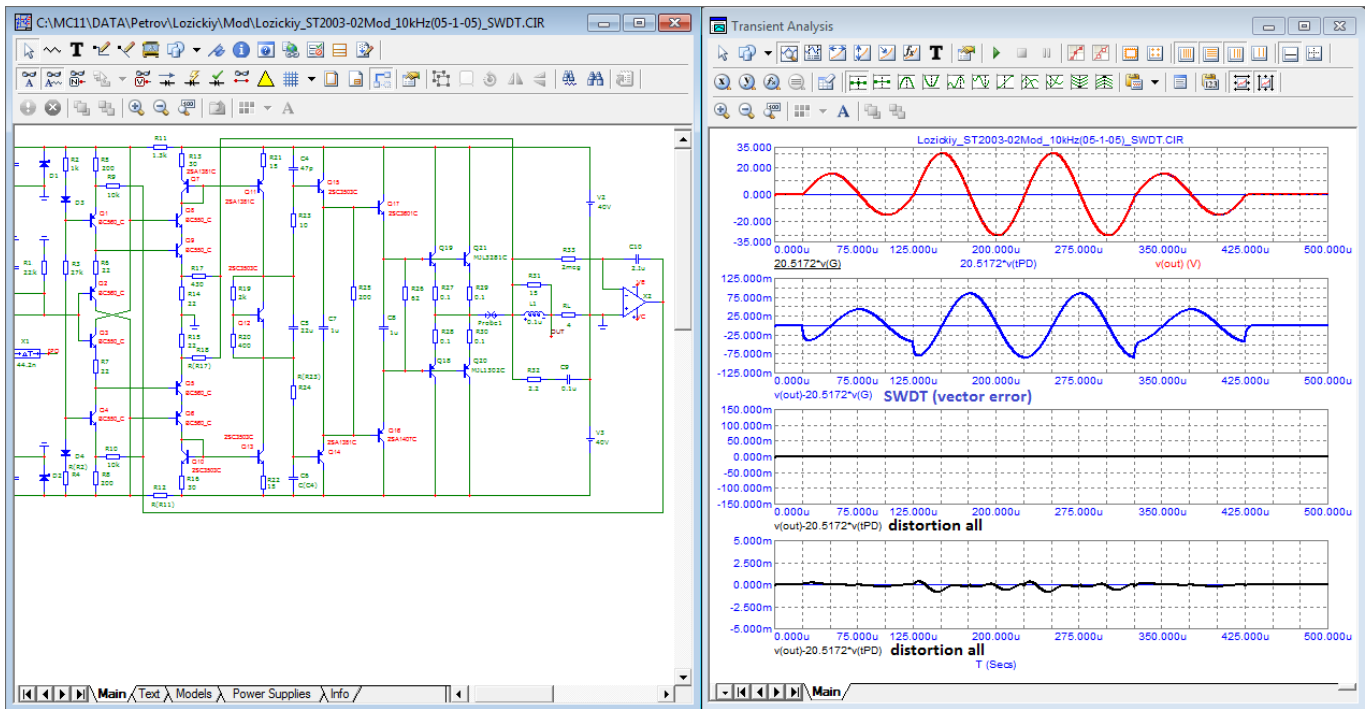


Fig. E. Amp Model [31] test result with 10 kHz bursts of different amplitudes (no phase break).

The graphs reflecting the distortion products (third and fourth) are shown on the same scale as in the test of the previous amplifier. The quiescent currents of the output transistors in both models are the same, and the output stages are approximately the same (in both OPS models - the “Darlington triple”), but due to the low feedback delay, the second amplifier copes better with switching and other types of distortion. The loss of useful information is reduced to almost zero.

Let's additionally carry out a test with bursts having different frequencies, fig. F.

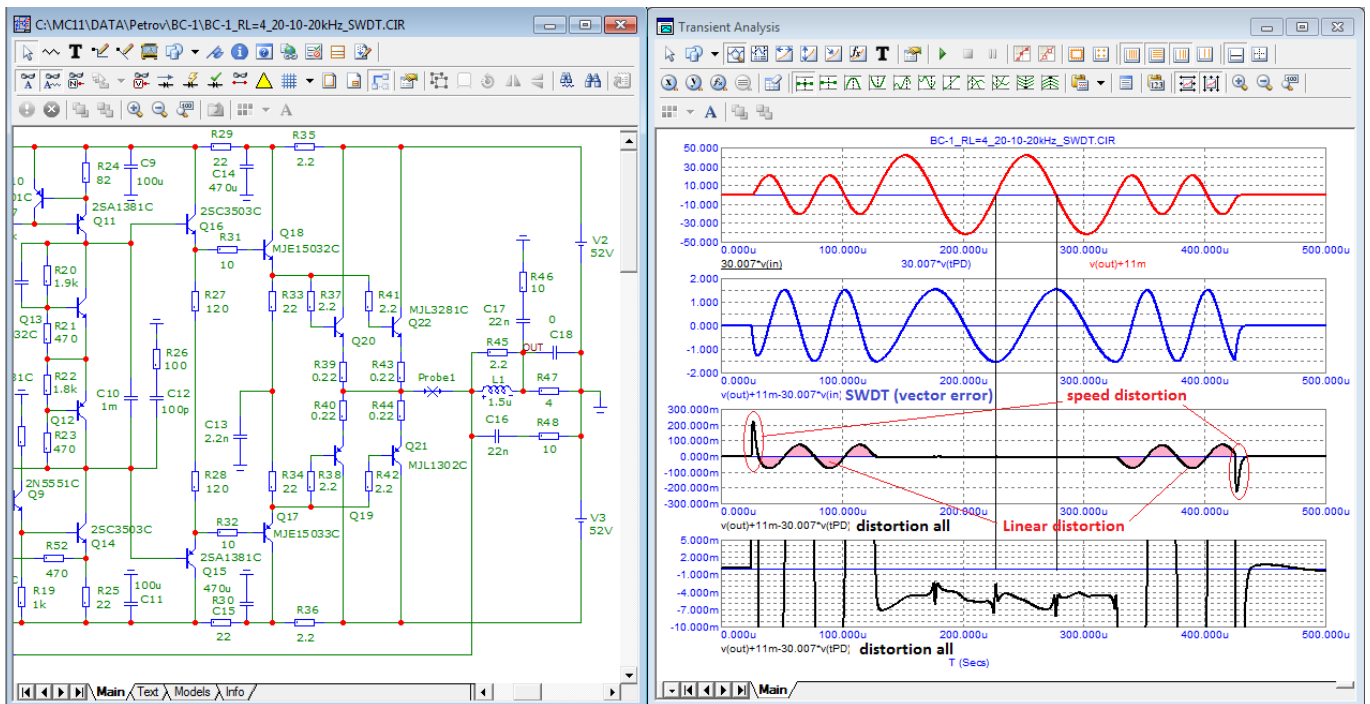


Fig. F. The test result of the BC-1 amplifier model using bursts of 10 and 20 kHz of different amplitudes (without phase break).

The test signal consists of successive bursts with a frequency of 20, 10 and 20 kHz. The third and fourth graphs show the distortion products. Despite the fact that the first burst at 20 kHz abruptly breaks off and the second burst at 10 kHz suddenly begins, there is no high-speed distortion at their junction. In the same way, there are no high-speed distortions at the junction of the second and third bursts. Velocity

distortion occurs only at the beginning of the first burst and at the end of the third.

Since the gain was selected for a frequency of 10 kHz, and at a frequency of 20 kHz there is a blockage in the frequency response with respect to a frequency of 10 kHz, the test additionally highlighted linear distortions for the first and third bursts (highlighted in pink).

Let's carry out a similar test of the second amplifier, fig. G.

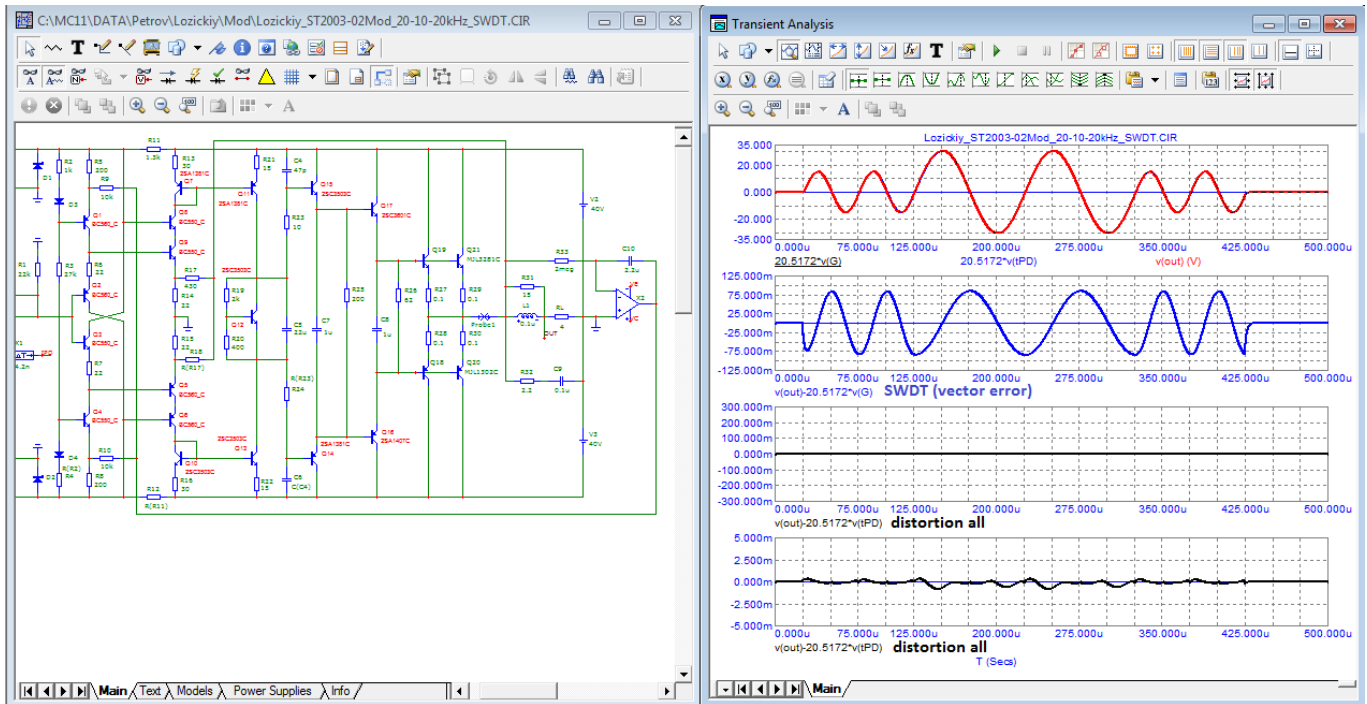


Fig. G. The test result of the amplifier model [30] using bursts of 10 and 20 kHz of different amplitudes (without phase break).

Here, as in the previous test of this model, all types of distortion are negligible (there are no speed or linear distortions).

Note.

It is advisable to use the compensation method for measuring distortions at the design stage of an amplifiers. No less successfully can this method be used to evaluate a previously developed amplifier without using computer simulation methods.

When measuring distortion with the compensation method, it is important to accurately measure both the delay time of the signal at the frequency under test and the gain. This greatly affects the measurement error.

For a more visual representation of the loss of the useful signal, we will conduct two amplifier tests [32] in which the delay time is 1.25 μ s. The amplifier is made according to the typical scheme of Douglas Self with the replacement of the “two” BJT with field-effect transistors. Unlike the typical circuit, feedback on the output current is added to adjust the output impedance.

To begin with, let's measure the loss of a signal with a frequency of 10 kHz, fig. H

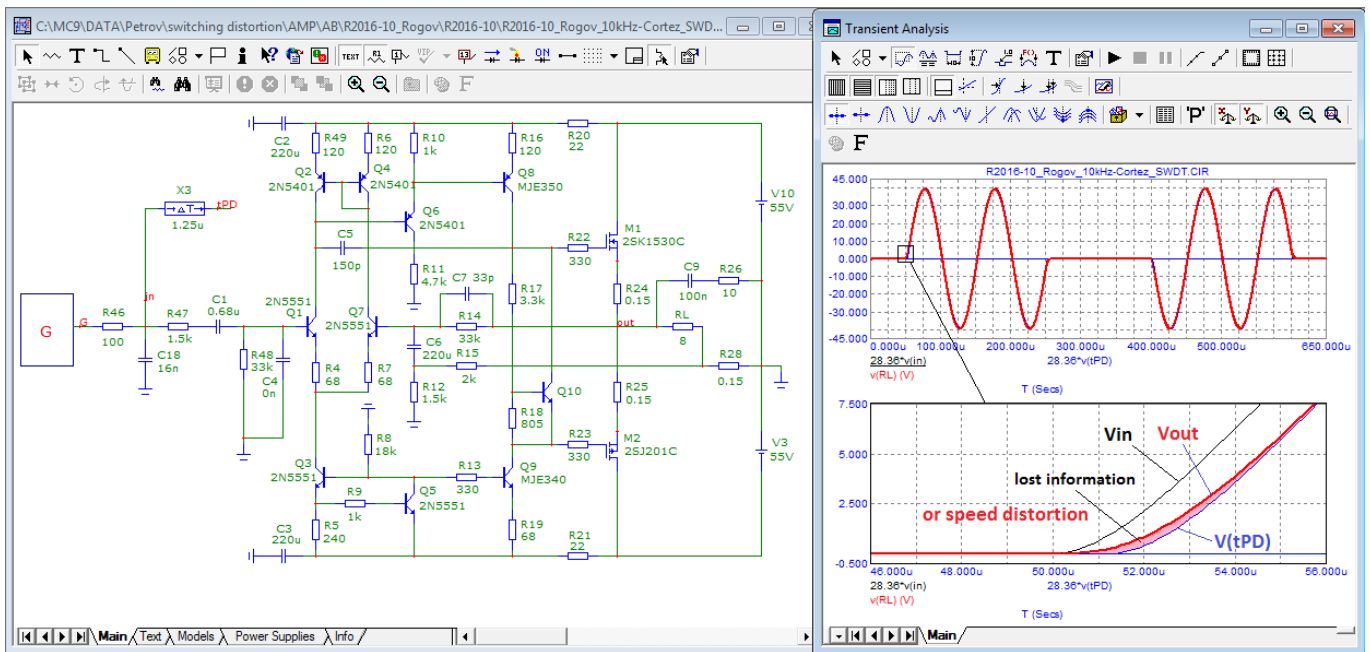


Fig. H. Loss of signal at the beginning of the burst

As a test signal, a 10 kHz signal was used in the form of bursts passed through a 100 kHz low-pass filter. The initial section of the burst is shown in a stretched form, on which the losses of the useful signal are highlighted in pink. The amplitude of losses reaches 500 mV, and their duration is several microseconds.

Let's carry out a similar test on a triangular signal, fig. I

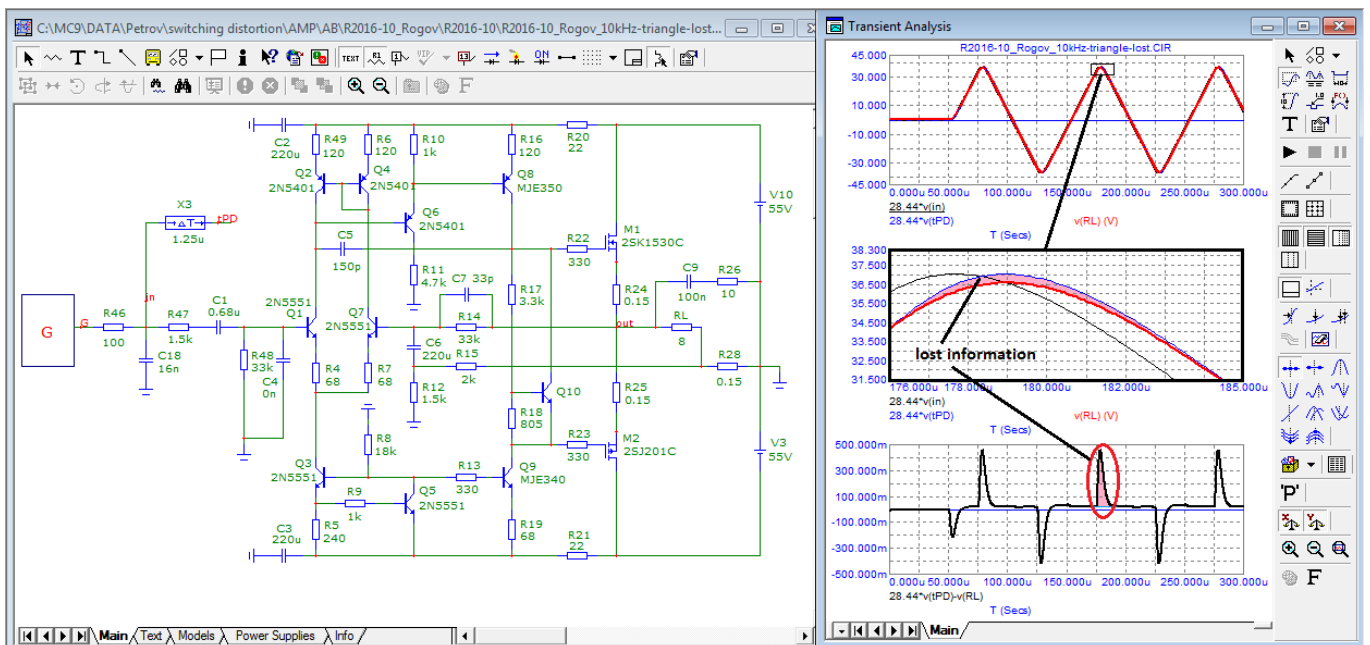


Fig. I. Signal loss at the vertices of the triangular signal (on voltage turns).

The top graph shows all three signals: the output signal (red) and the normalized input signal and tPD signal. Visually, all three signals practically merge. The second graph shows the top of the signal in a stretched form where all three signals are visible. The loss area is highlighted in pink, which in amplitude reaches 500 mV, and in duration is about 10 μ s. The third graph shows the result of subtracting the output voltage from the delayed input voltage. Here we also see the loss of the useful signal at each turn of the output voltage. What is 500 mV at an output voltage amplitude of 30 V. This is no less than 1.5% loss of information. Moreover, these losses are in no way related to the amplitude of the useful signal. That is, the same 500 mV can be lost even with a signal amplitude of only a few Volts, and this will no longer be 1.5%, but tens of percent of the deviation of the output voltage from the expected one.

Let me remind you once again that this kind of loss of a useful signal is not detected by any other types of testing except for the compensation method.

During operation, all amplifiers introduce various kinds of distortion. There are linear distortions (a change in the amplitude of the signals and their phase without adding harmonic components, which usually occurs after the end of transients, i.e. in steady state) and non-linear distortions caused by non-ideality of both active and passive elements. And although linear distortions (deterministic distortions due to bandwidth limitation and uneven amplitude-frequency response) do not depend on the signal level and do not add new spectral components, they change the amplitude and phase relationships (due to the unequal propagation speed of the spectral components) between the individual signal components and thereby distort the temporal structure, i.e. introduce distortions in the time domain and thereby change the shape of the signal, affecting the timbre of the sound.

The coefficient of non-linear distortion was first measured by K.Kupfmüller [1] and called them the clear factor (chatter factor). By the beginning of the 1950s, more than 20 different methods for measuring distortions had already been proposed.

According to the properties of the test signal and as their correlation with subjective quality assessments (SQA) grows, V. Rakovsky divided all these methods into five main groups [2]:

- 1) single tone method (method for measuring total harmonic distortion - THD (Total Harmonic Distortion));
- 2) methods of two tones (intermodulation distortion - IMD);
- 3) methods with a discrete spectrum (multitone);
- 4) continuous spectrum methods (white noise, pink noise, pseudo noise);
- 5) methods with a working signal (Sapozhkov's compensation method). Later, this is the Baksandall-Akulnichev method. Looking ahead, let's add another group here:
- 6) methods of working with real Hafler-Carver audio signals.

Each of these groups has varieties that differ in the methods of recording nonlinearity, isolating distortion products, etc.

Matti Ojala devoted a lot of time to testing methods, his research formed the basis of the DIM-30 and DIM-100 standards.

Developers such as John Curl pay special attention to the frequency of the first pole and higher harmonics from the 7th. Since negative feedback mainly suppresses the lower harmonics (2nd and 3rd) and almost does not affect the higher ones, John Kerl tries to get by with a minimum depth of feedback sufficient to bring the traditionally measured parameters to the minimum required by standards for Hi-Fi category amplifiers.

On the other hand, as practice shows, the OOS depth of less than 30 dB is ineffective.

Lynn Olson in [3] writes:

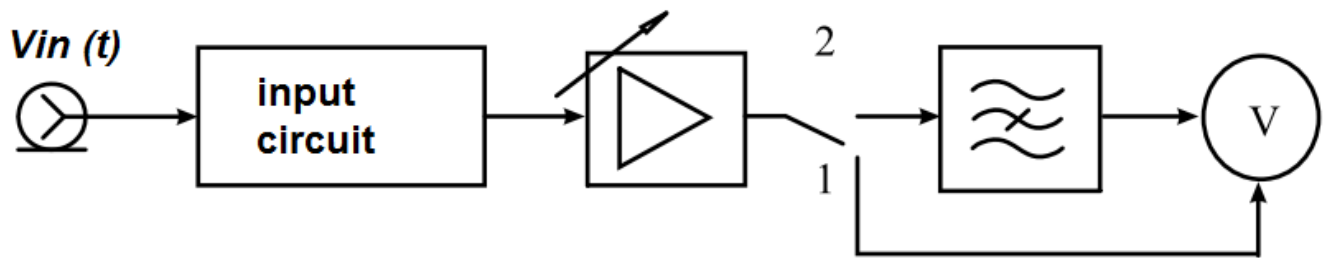
“In electronics, the correlation between distortion percentages and subjective assessment of sound is almost zero. The lower order harmonics are almost inaudible compared to the higher ones, although they dominate the THD measurement figures! The arrow of the meter tells the developer complete nonsense.

It's time to take a break from the myth of "euphonious distortion" and find those real subtle sources that create distortion in the amplifier that the human ear detects. Once we find measurement methods that actually help, rather than hide the truth from us, it will become easier to develop listener-friendly electronics.”

Unfortunately, to this day, the first method of testing is the main one. It's like in a well-known joke: at night we are looking for lost car keys under a lamp, and not where they were lost - because it is lighter there.

Using a single tone, bridge-based meters were used to suppress the main signal and highlight all harmonics at once, as well as methods for separately measuring the voltage of each harmonic using spectrum analyzers.

However, THD meters based on notch filters are the most widely used. The block diagram of such a meter is shown in fig. 1.



block diagram of a nonlinear distortion meter

1 - calibration, 2 - measurement

Fig. 1 Block diagram of the harmonic distortion meter

Let's try to figure out what is the reason for the low rating of this kind of meters, that they are treated as "probes" to control the compliance of amplifiers with their specifications during the production process (in the technical process).

There are a large number of notch filters, both passive and active, based on op amps. In order to evaluate the contribution of distortions to the useful signal at a level below 0.001% (-100 dB), the level of fundamental harmonic suppression must be 10...20 dB deeper, i.e. not less than 110...120 dB. To begin with, consider the simplest passive filter of the 4th order at a frequency of 10 kHz based on LC elements, Fig. 2

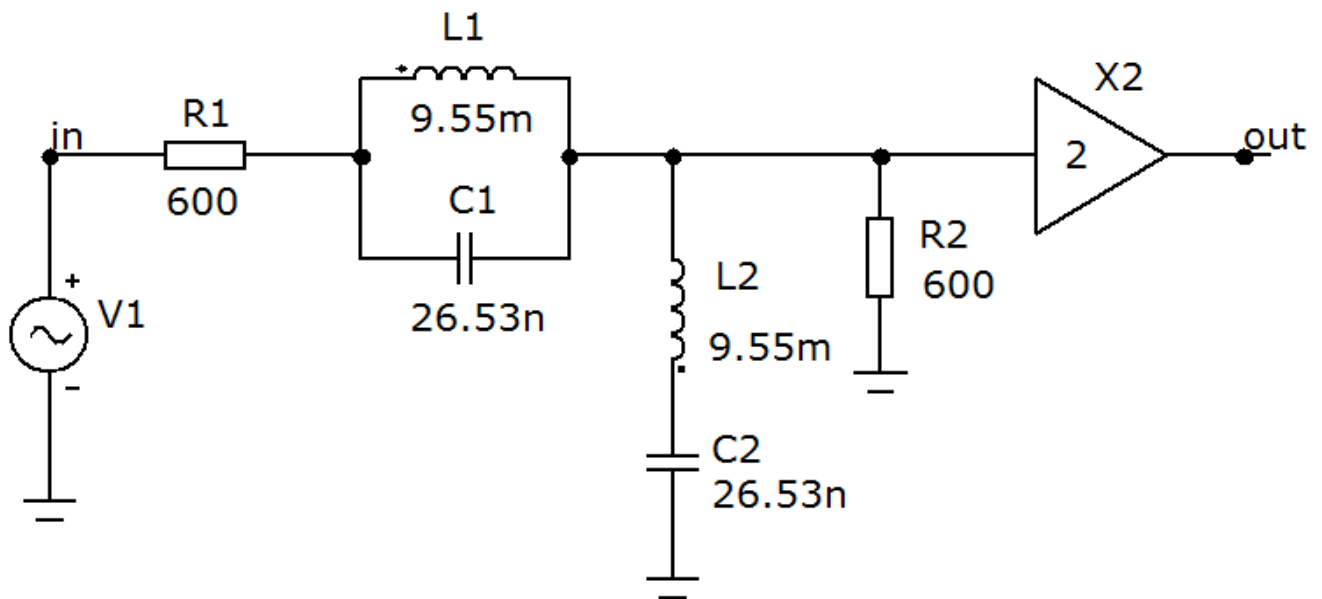


Fig. 2. 4th order Bessel notch filter

The frequency response of such a filter is shown in Figure 3

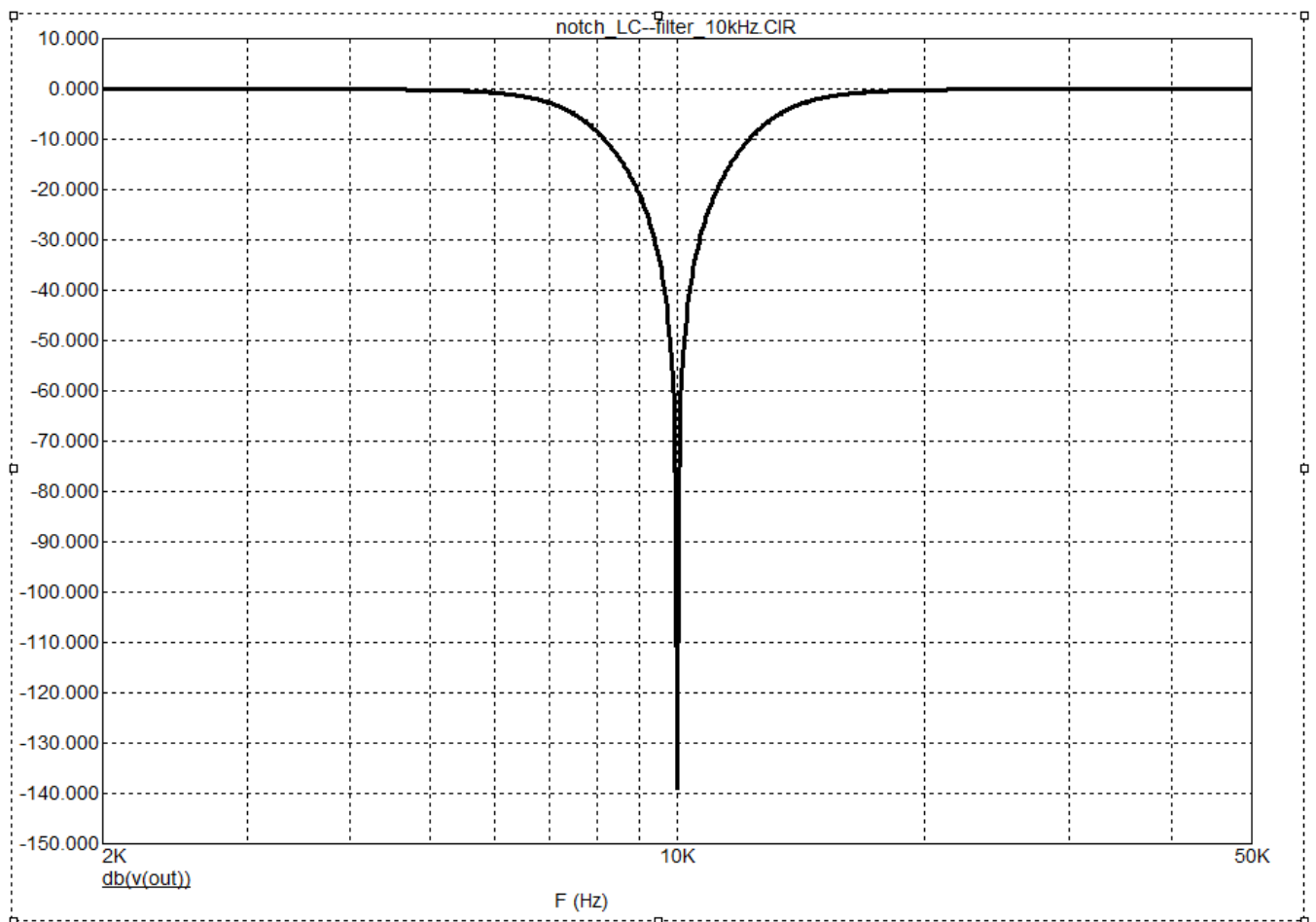


Fig. 3. The frequency response

As can be seen from the frequency response graph, the suppression of the fundamental harmonic is almost 140 dB, and the nearest 2nd harmonic and subsequent ones are transmitted without attenuation.

The next parameter that interests us is the settling time, i.e. time after which the filter can adequately measure distortion. To do this, we add to the fundamental tone with a frequency of 10 kHz and a level of 3 V (peak) a signal with a frequency of 30 kHz with a level of 3 μ V (-120 dB), fig. 4.

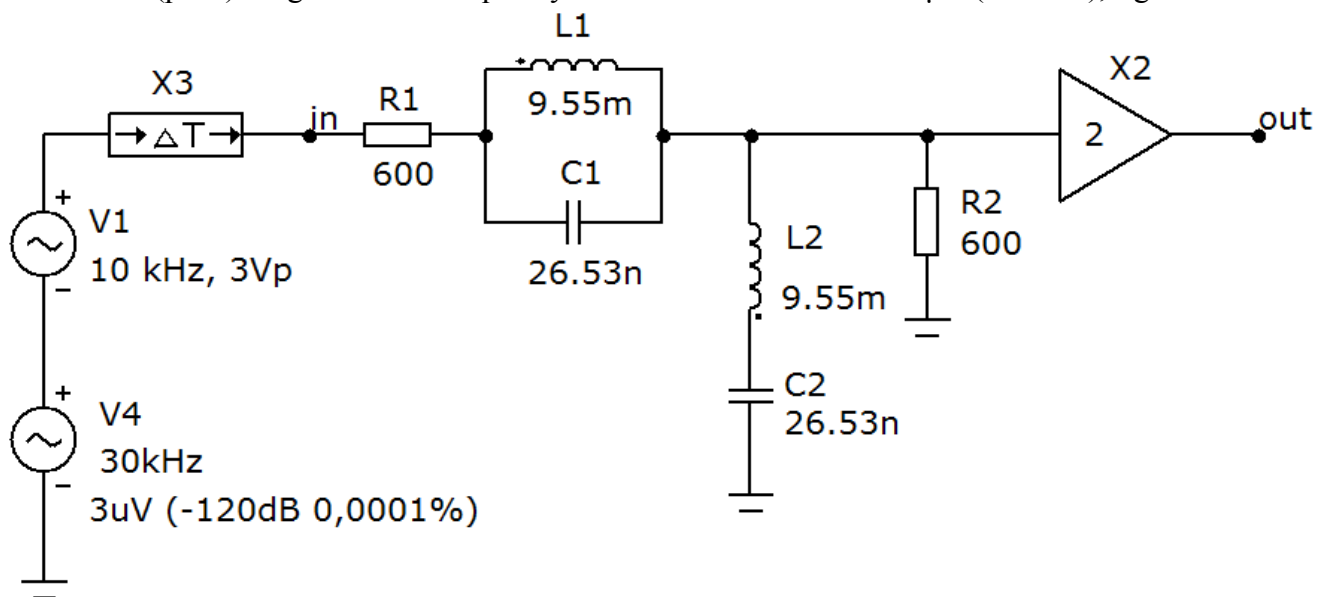


Fig. 4 Scheme for measuring the third harmonic of a signal with a frequency of 10 kHz

The third harmonic is simulated by an ideal V4 oscillator with an output frequency of 30 kHz. We start Transient/Analysis and look at the result, fig. 5

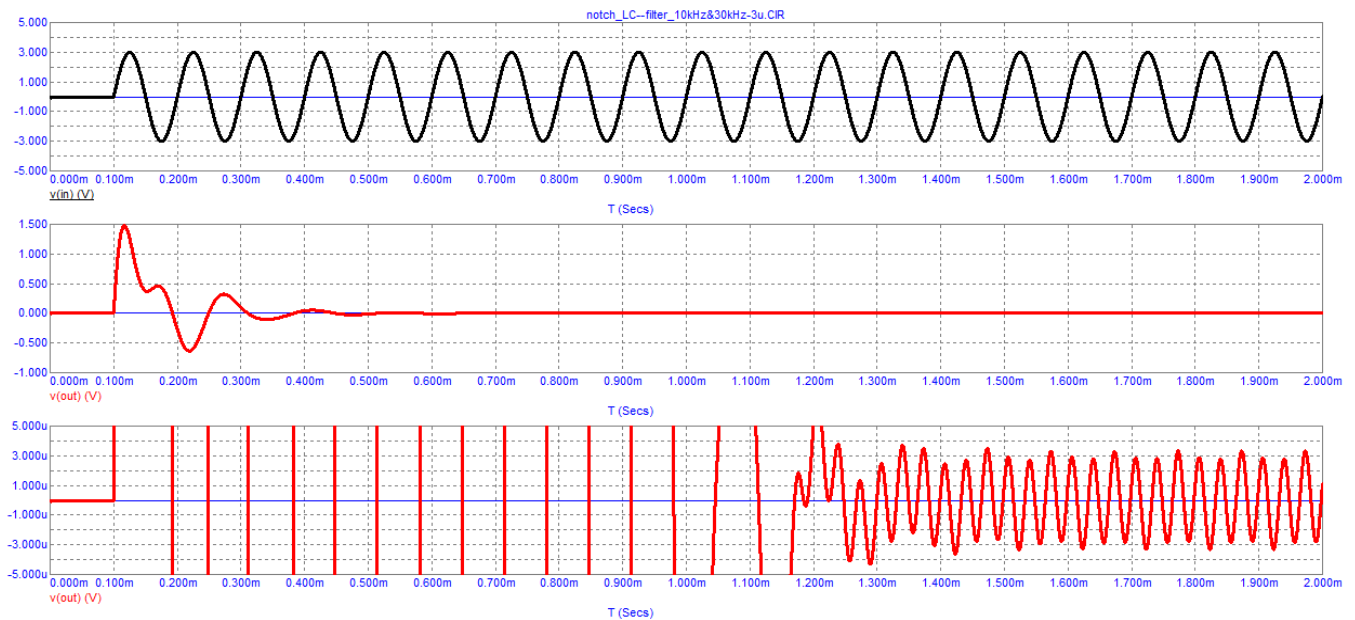


Fig. 5 3rd harmonic measurement result using the simulator

The test result shows that the transient distortion in the filter itself lasts about 1.5 ms, i.e. about 15 times longer than the period of the measured signal. Thus, it is not possible to measure distortions in the first periods (as it is possible to do with the help of the microcap program of the 9th version), and it is in the first periods (the largest contribution in the first) that the main distortions occur in the amplifiers due to transients, which erroneously referred to as linear distortion. Let me remind you that linear distortion occurs at the end of transients, i.e., in steady state.

Of the active notch filters, there are options based on a double T-bridge, a Wien-Robinson bridge, etc. But the suppression depth of such filters barely reaches 40 ... 50 dB, which is clearly not enough. Consider a notch filter based on phase shifters, [4], fig. 6

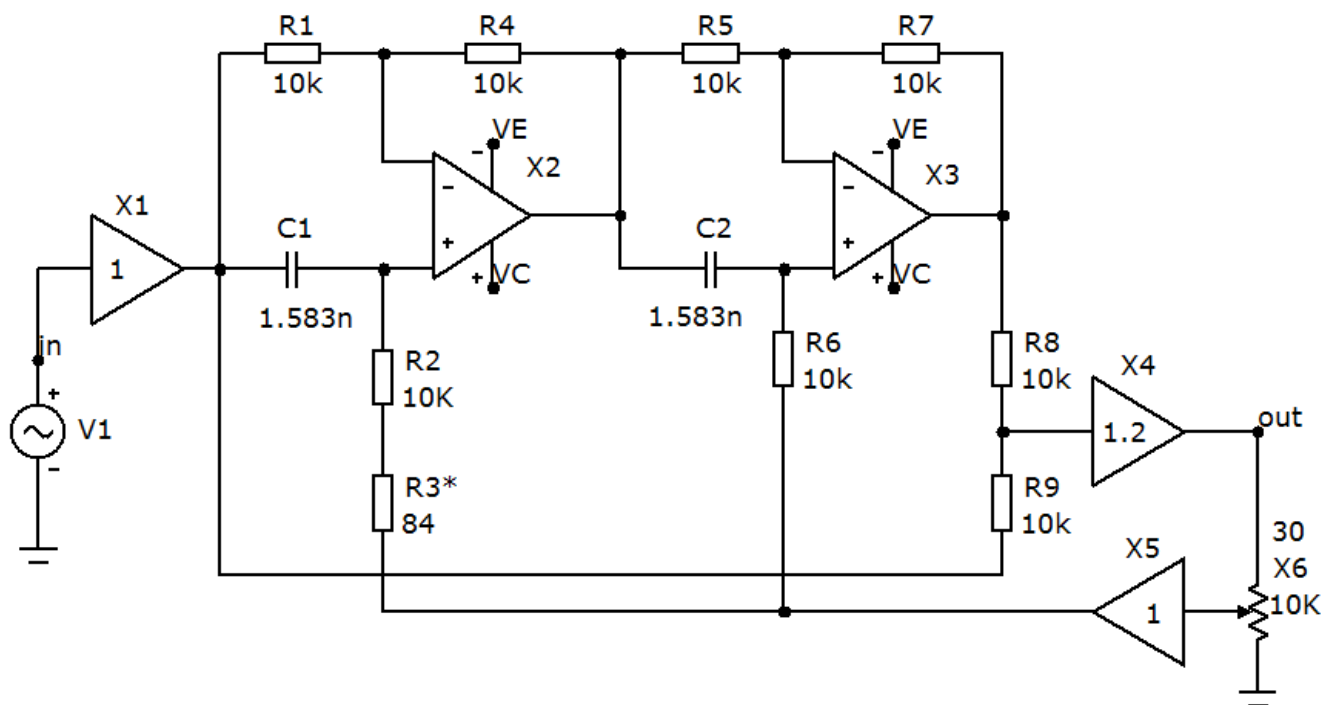


Fig. 6. Notch filter circuit on phase shifters

Such a filter suppresses the fundamental tone with a frequency of 10 kHz by more than 95 dB. To increase the degree of suppression, we turn on two such filters in series and measure the frequency response, fig. 7

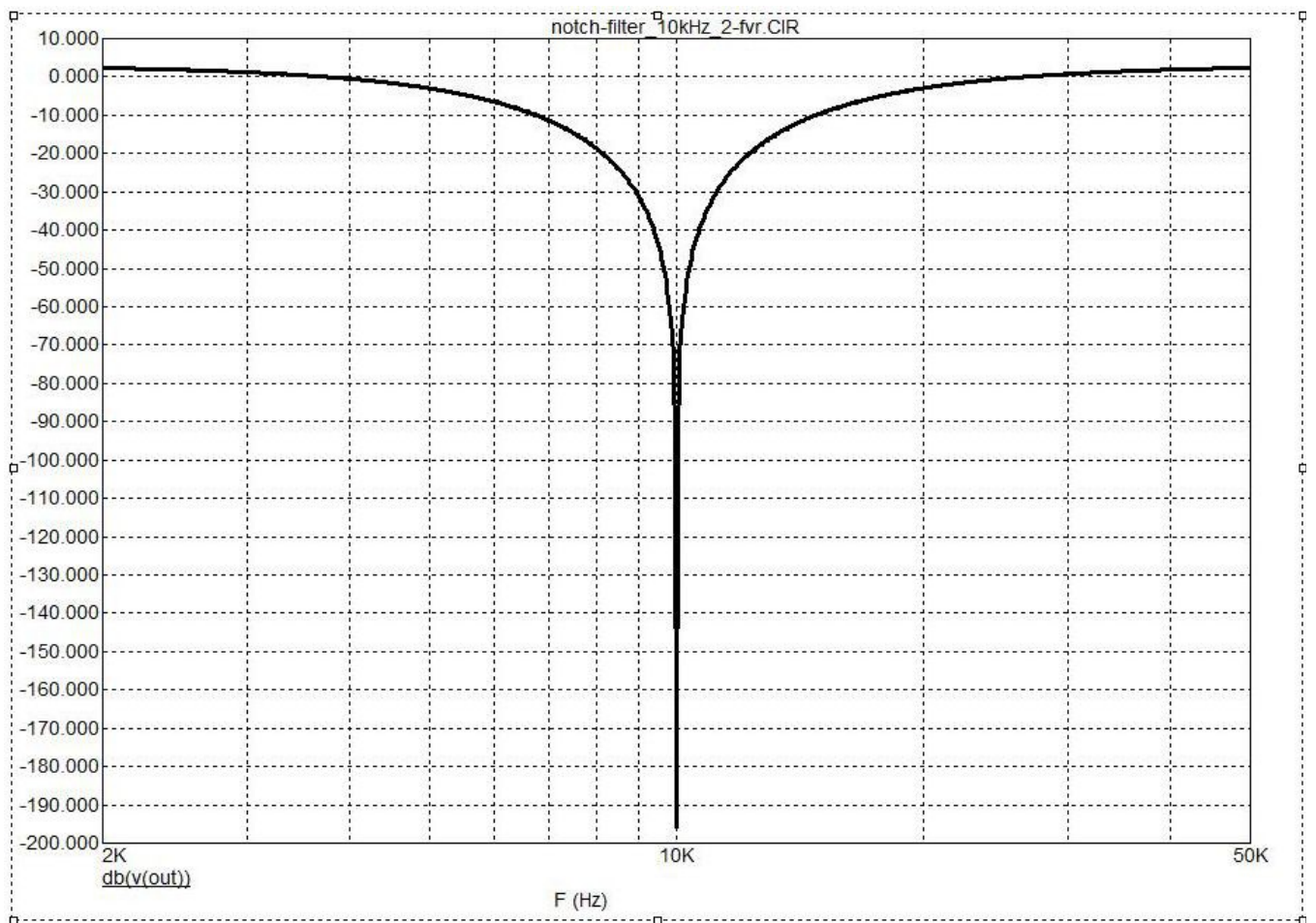


Fig. 7. frequency response of 2 notch filters connected in series on phase shifters

Let's add a signal with a frequency of 30 kHz with a level of 30 μ V (-100 dB) to the test signal, and check the settling time, fig. 8

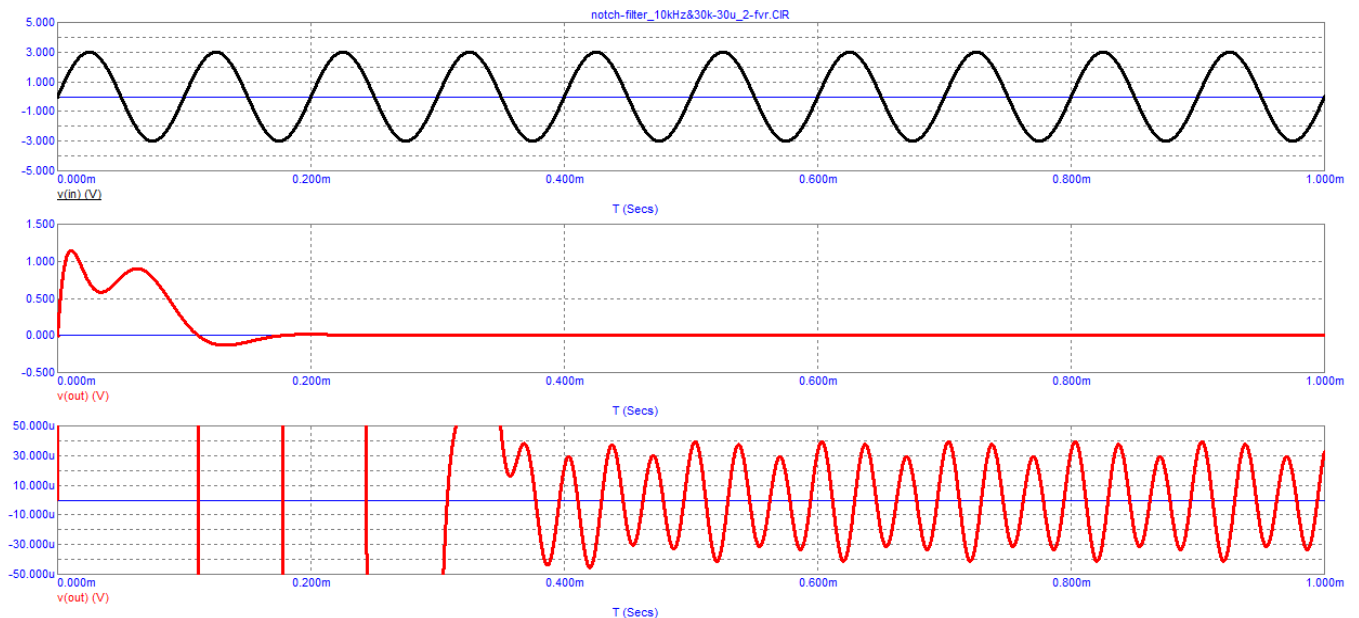


Fig. 8. 3rd harmonic measurement result

The test shows that the settling time of such a filter is almost 3 times shorter. However, despite almost 50dB more rejection than the previous filter, the filter is unable to adequately measure distortion even as low as 0.001% (-100dB) due to the distortion introduced by active elements (op-amps). Models of high-speed precision op amps of the OPA627 type were used as the op amp. Thus, even with this filter, at least 4 pitch periods must be skipped. By the way, in the latest versions of the microcap (11th and 12th) it

is impossible to set less than 4 periods to calculate the amplifier for distortion (harmonic spectrum, THD).

The sinusoidal voltage is generated by a constant amplitude vector rotating at a constant angular velocity.

This explains why the test results of this method do not correlate with the sound quality. After all, the main distortions occur at the moments of change in dV / dt (at the moments of change in the voltage vector or the angular velocity of its rotation), which takes place continuously in real sound signals.

The situation is somewhat better with two-tone and multi-tone methods, however, due to the fact that there are also simple fixed signals in the spectrum, the correlation with sound quality that was expected does not occur.

The spectrum of a signal is the result of decomposing a signal into simpler orthogonal functions in the basis. As a decomposition, the Fourier transform, expansion in Walsh functions, wavelet transform, etc. are usually used. In simulators, the Fourier transform is mainly used.

At the beginning of 1950, W. Wolf applied for a patent “Method for measuring non-linear distortions” [5], fig. 9.

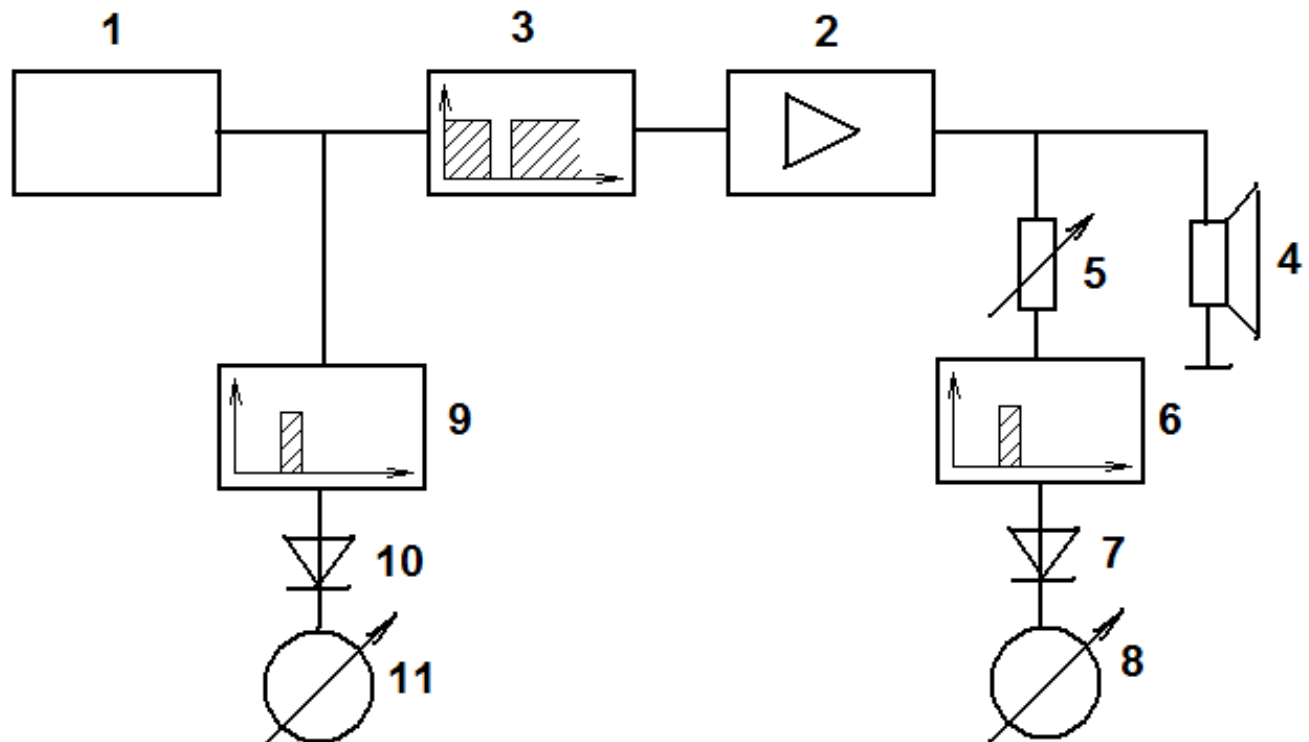


Fig. 9. Block diagram of the NI meter according to patent SU90185

The essence of the proposed measurement method is illustrated by the diagram. From a continuous spectrum sound generator of equal amplitudes 1, the voltage is fed to the input of the investigated path 2 through a notch filter 3, which delays a narrow frequency band from f_1 to f_2 . From the output of path 2, the voltage is supplied to the load 4 and in parallel through the voltage regulator 5 to the bandpass filter 6 with a bandwidth from f_1 to f_2 . From filter 6, voltage is supplied to detector 7 and then to galvanometer 8.

Since the passband of filter 6 is equal to the stopband of filter 3 (or somewhat worse than it), then at the output of the studied path, in the case of its linearity, there should be no currents whose frequencies lie in the passband of filter 6, and galvanometer 8 should not deviate. The deviation of the galvanometer needle will indicate the appearance at the output of the studied path 2 currents in the frequency band $f_1 - f_2$, i.e. frequency currents, which, due to the presence of filter 3, were absent at the input of the path.

These generation of currents in the frequency band $f_1 - f_2$ are the result of the appearance of harmonics from the lower frequency components contained in the input signal, and the result of the formation of combination frequencies falling on the band $f_1 - f_2$.

The deviation of the arrow of the galvanometer 8 is proportional not only to the degree of nonlinearity of the studied path, but also to the amplitude of the voltage acting at its input. To separately account for both of these factors, another band-pass filter 9, detector 10 and galvanometer 11 are introduced into the circuit with the same characteristics as those of similar elements 6, 7, 8.

With the help of voltage regulator 5, an attenuation equal to the gain of path 2 is introduced into the filter circuit 6. In this case, the deviation of the galvanometer 11 will be proportional to the emf. useful signal in the frequency band $f_1 - f_2$, and galvanometer 8 - in proportion to emf. distortion, newly formed by the path 2 itself in the same frequency band. The ratio of the readings of the galvanometer 8 to the readings of the galvanometer 11, with an appropriately selected attenuation of the voltage regulator 5, will be a measure of the non-linear distortions created by the investigated path.

For the convenience of reading galvanometers 8 and 11 can be replaced by a ratiometer, the angle of deviation of the needle of which is proportional to the ratio of the currents flowing through its coils.

Since sound energy is unevenly distributed over the frequency spectrum, it may turn out that the distortion measured in the frequency band $f_1 - f_2$ is not characteristic of other parts of the spectrum. Therefore, filters 3, 6 and 9 should be made tunable over the range of audio frequencies while maintaining the same width of their bandwidth (and, accordingly, delay). With a sufficiently fast tuning of the mentioned filters, the device will give an idea of the average energy distortion value for a time equal to the filter tuning period. In this case, it makes sense to make detectors 7 and 10 according to the scheme of pulse meters with a correspondingly selected time constant.

The main reason for filing an application is that tests performed by applying a pure tone (or several pure tones) to the input of a quadripole do not take into account the spectral composition of a real sound signal. Therefore, the magnitude of those non-linear distortions that arise in the equipment when it is modulated by a voltage of a complex spectrum, characteristic of a real sound program, is far from being fully reflected. The possibility of occurrence of currents of combination frequencies, which sharply increase nonlinear distortions as a result of amplitude-phase conversion of signals, is also not taken into account. And most importantly, tests with pure tones cannot highlight information that is lost by amplifiers when amplifying real audio signals.

According to this method, at the AES-126 held in 2009 in Munich, a well-known specialist in the field of electroacoustics prof. A. Farina presented a report [6].

However, the Wolf method turned out to have a significant drawback - despite the possibility of tuning filters over the sound range, it is impossible to determine exactly in which frequency band the amplifier has the greatest nonlinearity.

This drawback was eliminated by a meter acting exactly the opposite [7]: 1/3 octave noise (white noise, pink noise or pseudo noise) is fed to the amplifier, and it is cut out at the output. As a result, distortion products generated by the noise signal remain in the spectrum, Fig. 10. And since the noise signal is modulated both in frequency and in amplitude, it just provokes distortions characteristic of transient processes in the first periods of tone signals.

Assuming that the high-frequency part of the audio range is most critical to the non-linear distortion of amplifiers, we can limit ourselves to several 1/3 octave bands starting from a frequency of 10 kHz. As test signals, you can use ready-made signals available on test disks to check the frequency response of the audio path.

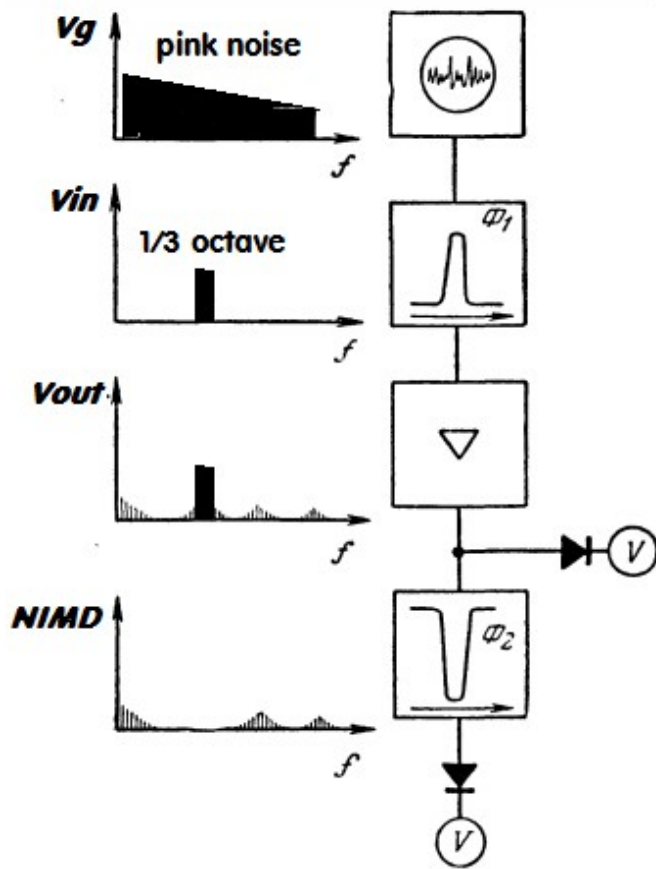


Fig. 10. Principle of operation of the INI using narrowband noise

More details about the results of comparative tests can be found in [8].

The next step in the development of measuring technology was the direct subtraction of signals after scaling and phase matching of the input signal to the output phase, [9]. This method (the method of cross-correlation between input and output signals) was defended by Doctor of Technical Sciences M. Sapozhkov in 1954. The same method was also published in [10]. This method was later published in [11] and became known as the Baxandall method. In Russia, this method was popularized by I. Akulinichev, who published a number of vector distortion indicators in the Radio magazine.

All types of distortions are most fully reflected in [12], and are illustrated in Fig. 11 and 12.

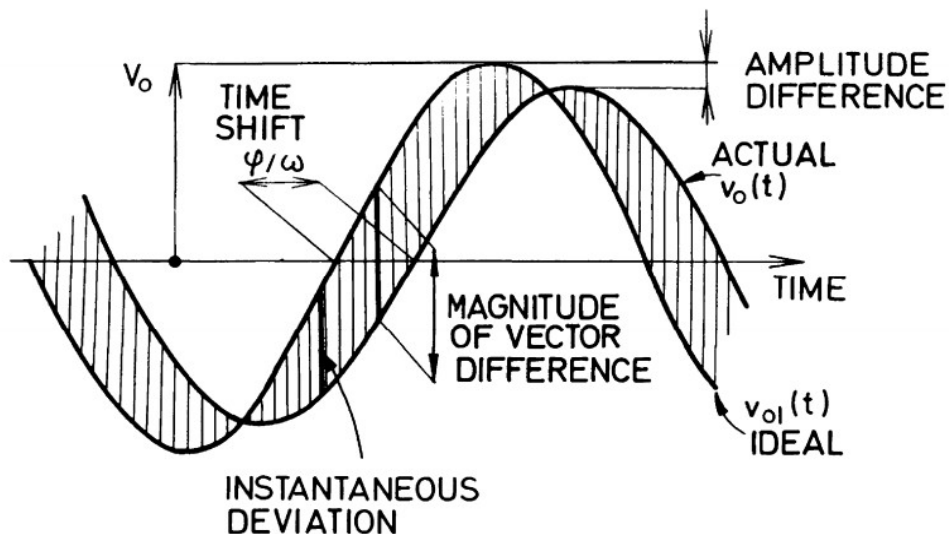


Fig. 11. Illustration of the relationship between vector, amplitude and phase errors in steady state when representing the signal as a function of time.

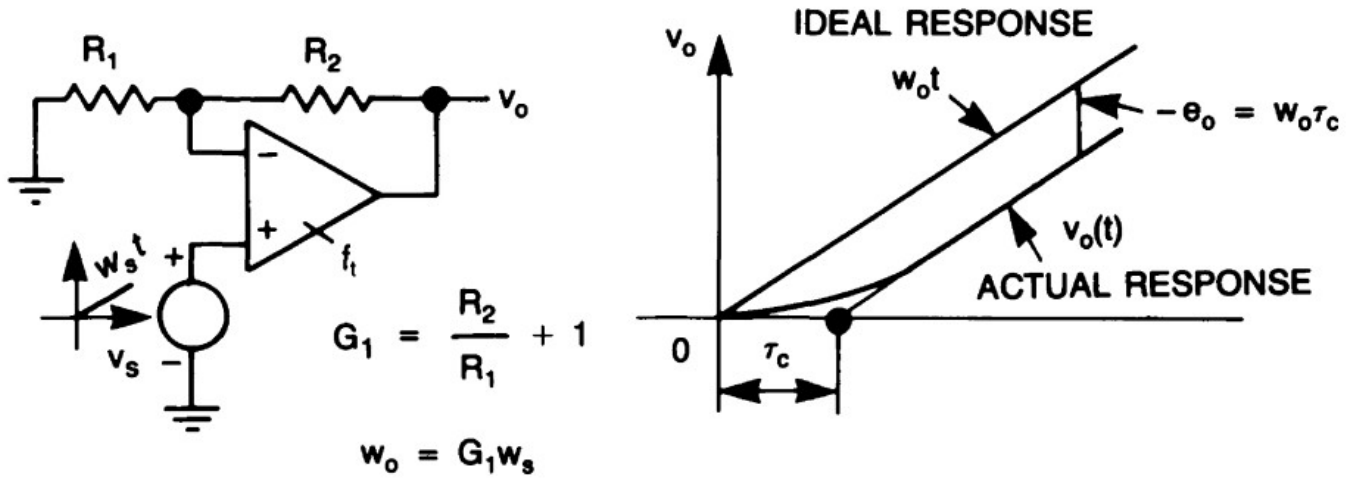


Fig. 12. The speed error of a non-inverting amplifier when excited by its linearly increasing voltage.

The types of distortions are deciphered in paragraphs 13 and 14 [12], we will give them in full with minor additions

13. The speed error of a first-order resistive operating circuit with a corner frequency $f_c = 1/2\pi\tau_c$, excited by a linearly changing signal - such that the signal at the output changes with the speed ω_{out} , is equal to:

$$\epsilon_{\text{blx}} = -\omega_{\text{blx}}\tau_c \quad (1)$$

where $\omega_{\text{blx}} = 2\pi f_{\omega} U_{\text{blx}}$ (2) (look. 14)

$\tau_c = RC$ and is physically equal to the signal delay (time Propagation Delay – tPD) in the passband

14. Both the speed error ϵ_{out} with non-harmonic and the vector error ϵ_v with harmonic action express the same basic dynamic limitation, associated with the presence of the interface frequency of the operating circuit f_c . In a first-order resistive operating circuit, these two errors are related by:

$$-\epsilon_{\text{blx}}/U_{\text{blx}} = \epsilon_v(f_{\omega}) \quad (3)$$

Substituting expression (2) into (1) we get:

$$\epsilon_{\text{blx}} = -\omega_{\text{blx}}\tau_c = -2\pi f_{\omega} U_{\text{blx}}\tau_c = -2\pi f_{\omega} U_{\text{blx}}(tPD) \quad (4)$$

Replacing the frequency with a period for the harmonic effect, we get:

$$\epsilon_{\text{blx}} = -2\pi U_{\text{blx}}\tau_c/T = -2\pi U_{\text{blx}}(tPD)/T \quad (5)$$

From paragraph 14 it follows that the speed error for non-harmonic effects and the vector error for harmonic effects are calculated using the same formula, which is not surprising, since the initial section of the sinusoid (more precisely, the initial section of the burst) can be considered with good approximation as a non-harmonic effect.

Methods for measuring the vector error consist in direct subtraction of the output signal from the input signal reduced in terms of level to the output signal (normalized voltage). For example, for an inverting amplifier, it looks like this, fig. 13 [12].

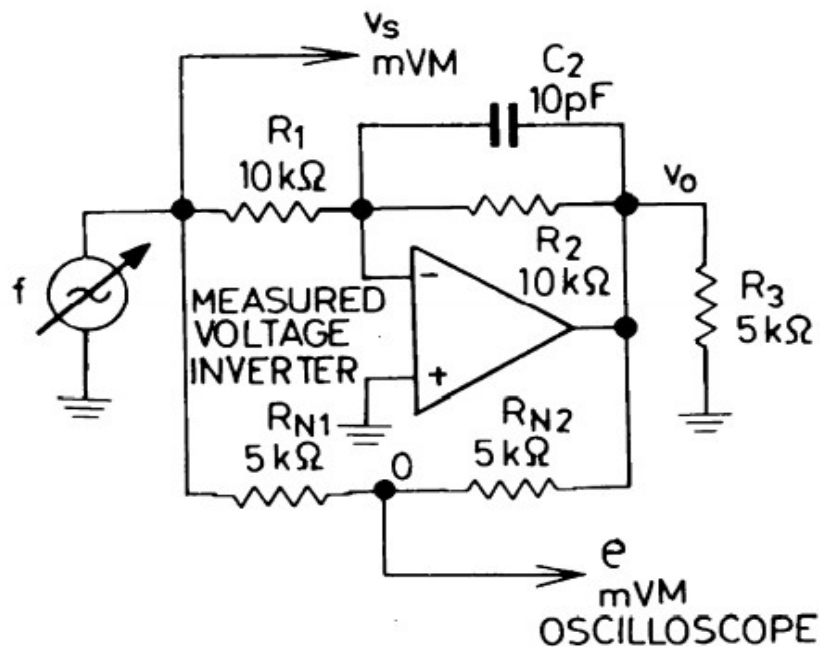


Fig. 13. Scheme for measuring the vector error of the voltage inverter

Since the gain is minus 1, the result of the subtraction is measured at the connection point of resistors of equal value.

Thus, according to Jiri Dostal, the vector error and the velocity error are essentially the same for sinusoidal signals in steady state. However, the initial section of the linearly increasing voltage, or, which is the same, the initial section of the first sinusoid of the burst (a small "triangle" in Fig. 12) was omitted from consideration. It is these distortions that I called speed distortions. This kind of distortion is related to the rotation of the voltage and would be better called "rotation distortion". This type of distortion occurs every time the signal deviates from the sinusoid (at the moments of change in both the signal frequency and its amplitude).

The developers needed a test that would correlate with sound quality. and David Hafler proposed such a test called SWDT ("straight-wire" differential test) [13], fig. 14.

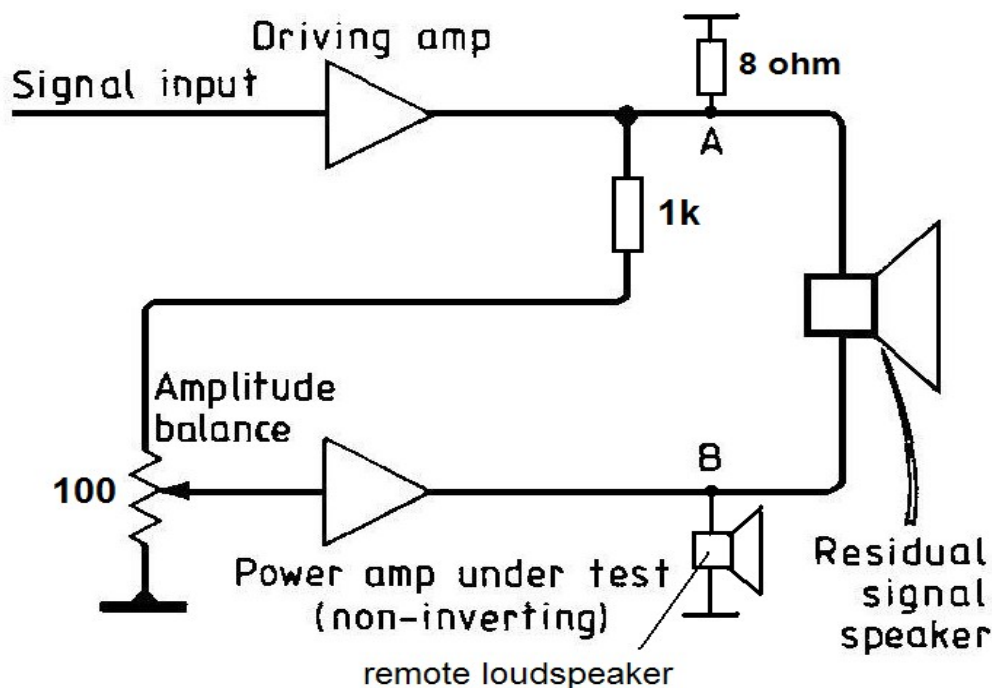


Fig.4. Hafler 'straight-wire' differential test.

Fig. 14. Block diagram of the Hafler vector error meter

The Hafler test consists in comparing the output voltage with a scaled (normalized) input voltage (with the input voltage leveled to the output when viewed from the input of the amplifier under test to point A through the attenuator). Obviously, if the input and output are identical, there will be no signal in the monitor speaker (headphones). Any sound heard after careful balancing will be distortion. For the "null test" Hafler suggested a level of -70 dB for the middle frequencies (signal attenuation by 3000 times or up to 0.03%) and -60 dB (signal attenuation by 1000 times or up to 0.1%) for higher frequencies of the audio range.

In one of the answers to questions on the article [13] on the application of the SWDT test, Hafler writes: "This is not a new idea. However, in the past, amplifiers have performed very poorly in this test, which has revealed all the shortcomings. The XL-280 may be the first amplifier to offer good zero signal with SWDT."

The idea is not really new, but audio amplifier designers want to see distortion products, not vector errors. But for this it is necessary to carry out a test taking into account the delay of the signal.

And although the Hafler test does not measure non-linear distortions, but only vector errors, nevertheless, in some cases it can be more informative than any other test.

Unlike the Sapozhkov-Baksandall method, this method does not use phase adjustment, which greatly simplifies the testing procedure, and on the other hand, it also takes into account distortions in the time domain, i.e. introduced phase distortions, which often lead to serious distortions associated with the amplitude-phase conversion of signals.

This idea was voiced in the report "Pitch, timbre, separation of sources and myths about sound reproduction through loudspeakers" by a well-known specialist in the field of psychoacoustics D. Griesinger at AES-132 held in Budapest in 2012. In his report, he showed that the phase relationships of the upper harmonics in the spectrum are extremely important, and also that when playing through loudspeakers, the structure of the amplitude and phase spectrum is always violated to one degree or another, so the sound is always poorer in timbre, clarity, clarity, etc. . than natural sound.

You may get the impression that the test result does not depend on the top amplifier in the circuit, because its task is only to give out the signal amplitude. In fact, it should be wide enough (high-speed) so as not to "blur" the subtle nuances of the sound material and not make it easier for the amplifier under test. Only in this case we will get more reliable information about the tested amplifier. If you use a "braked" amplifier (this may be one channel of the tested stereo amplifier used as a driver), then the tested amplifier may look better than it actually is.

Regarding the Hafler test, John Kerl said the following [14] (p. 40): "With all due respect to David Hafler. His amp may have passed his test, but this one the amplifier is not considered "perfect" by any standard and is usually "upgraded" amateurs, with marked "improvement". It must be assumed that he meant the XL-280 amplifier.

At first glance, this statement may seem paradoxical. In fact, a lot depends on the test signals. If pure sinusoidal signals are used, the measurements will differ little from the steady state THD measurement, with the ensuing consequences. The most optimal signal for such testing is a triangular signal with a frequency of 10 kHz passed through a 100 kHz low-pass filter. Such a signal is very convenient as a test signal in all respects, the rate of rise / fall of the signal fronts is close to the maximum rate of rise encountered in real audio signals, and at the top of the signal there is a turn of the signal - a change in voltage dV / dt characteristic of the subtle nuances of a real audio signal (reveals speed distortion twice per period). It was the test using such a signal that showed that the XL-280 amplifier model introduces high-speed distortion up to 1% or more instead of 0.1%.

When using a sinusoidal signal, equal amplitudes at the outputs of the amplifiers and no distortion, the difference signal will also be a sinusoid, but phase shifted, the amplitude of which depends on the signal propagation delay of the amplifier under test (tPD):

$$a \sim 2\pi A * tPD / T \quad (6)$$

where

T- is the signal period, μs ;

A - is the signal amplitude at the amplifier outputs, V

tPD – signal propagation delay time, μs

where

$$tPD = \sim aT / 2\pi A \quad (7)$$

From formula (6) it can be seen that the level of the vector error is inversely proportional to the period (directly proportional to the frequency). Therefore, at higher frequencies, it is most difficult to provide small vector errors.

Calculate the maximum allowable delay t_{PD} using formula (7) for a frequency of 20 kHz with a vector error of -60 dB, i.e. with $a/A = 0.001$

$$t_{PD} = \sim aT/2\pi A = 0,001 * 50000 \text{ нс} / 6,28 = 8 \text{ нс}$$

Thus, the signal propagation delay time should not exceed 8 ns. I must say that this is a rather stringent requirement for UMZCH, since the vast majority of amplifiers in operation have t_{PD} from 200 ... 300 ns to 1.5 μs or more.

If we focus on the amplitude accuracy of signal transmission of 0.01%, then, in accordance with the conclusions of [12], the bandwidth of the amplifier with single-pole correction should be at least:

$$f_6 = (100/\sqrt{2})f_1 \approx 71f_1 \quad (8)$$

We substitute the upper frequency of the audio range into formula (8), as a result we obtain the required bandwidth of 1.4 MHz. This means that the signal propagation delay time should not exceed:

$$t_{PD} \leq 1/2\pi f_B = 1/(6,28 \cdot 1400000) = 0,00000013 \text{ с} = 0,00013 \text{ мс} = 0,13 \text{ мкс} = 130 \text{ нс}$$

Thus, in accordance with the calculations by Jiri Dostal's formulas, the bandwidth of the amplifier must be at least 1.4 MHz, and the delay time of the signal in the amplifier itself (excluding the delay caused by the input RF low-pass filter) for high-quality amplifiers with OOS should not exceed 130 ns , but in that case it would not meet the more stringent requirements of the Hafler SWDT test.

In some modern top-end amplifiers, this parameter began to be indicated. For example, the technical specifications for the MIMESIS 9.2 amplifier indicate:

GROUP DELAY

- Propagation Delay: < 100 ns stable from DC to 200 kHz

A similar "null test" was also used by Bob Carver [15], fig. 15

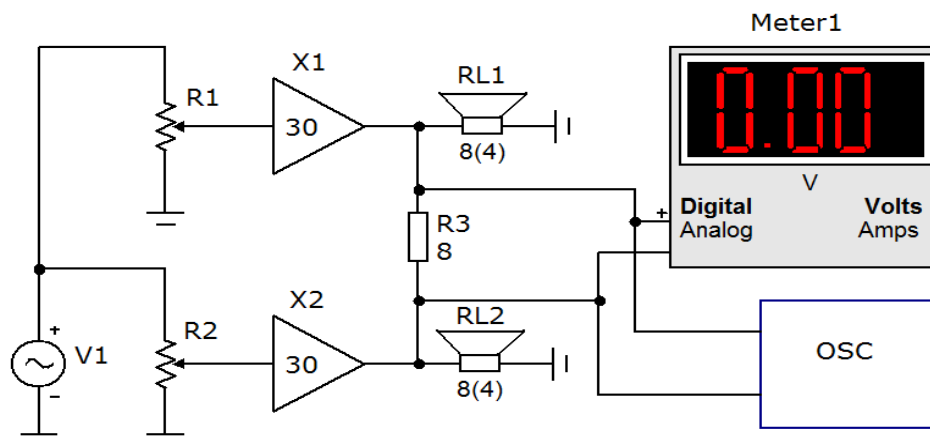


Fig. 15. Block diagram for measuring the relative level of distortion with respect to the standard

Unlike the original Hafler test, the Carver test uses a comparison between the output of the "reference" X_1 and the output of the amplifier under test, X_2 . As the experiment showed, adjusting the parameters of an amplifier of the middle price category to the parameters of an expensive "standard" with an accuracy of -70 dB (0.03%) ensured the identity of their sound.

Since we are striving for perfect accuracy, when modeling with simulators, an ideal amplifier with a gain equal to the gain of the amplifier under test at the signal frequency can be used as a reference amplifier. Then, to measure the introduced distortion by the compensation method, it is enough to delay the output signal of an ideal amplifier using an ideal delay line for the signal propagation delay time (t_{PD})

in the amplifier under test, and all that remains after subtraction is the distortion introduced by the amplifier.

The Carver method can also be used to measure the gain in distortion at maximum power (or any other) for example, in relation to an output power of 1 watt. To do this, amplifiers X1 and X2 must have the same parameters (for example, monoblocks of the same brand). At the output of one of them, we turn on the divider in parallel with the load (we take the resistor that is connected to the “common” equal to the load, and calculate the upper resistor so that at the connection point of the divider resistors there is a voltage corresponding to an output power of 1 W). It remains in the amplifier with a divider parallel to the load to set the output power close to the maximum, and in the second amplifier, set the output voltage equal to the voltage at the division point of the resistors of the first amplifier with the input attenuator. If the parameters of the amplifiers are the same (it is important that they are equal to tPD), then vector distortions will be compensated and we will get a pure increase in distortions in the bridge diagonal. They can be evaluated using an oscilloscope, FIR or a spectrum analyzer.

It is known from psychoacoustics that human hearing is most sensitive to the rate of phase change, i.e. to the group delay (GDT) and its variations in the frequency domain. In fact, the constancy of the group delay is important not only in the audio range, but also far beyond it (up to 1 MHz and beyond) and it is desirable that the effect of the load on the group delay is as small as possible.

The relationship of vector errors with the delay time is easy to follow in the following figures, 16 - 18.

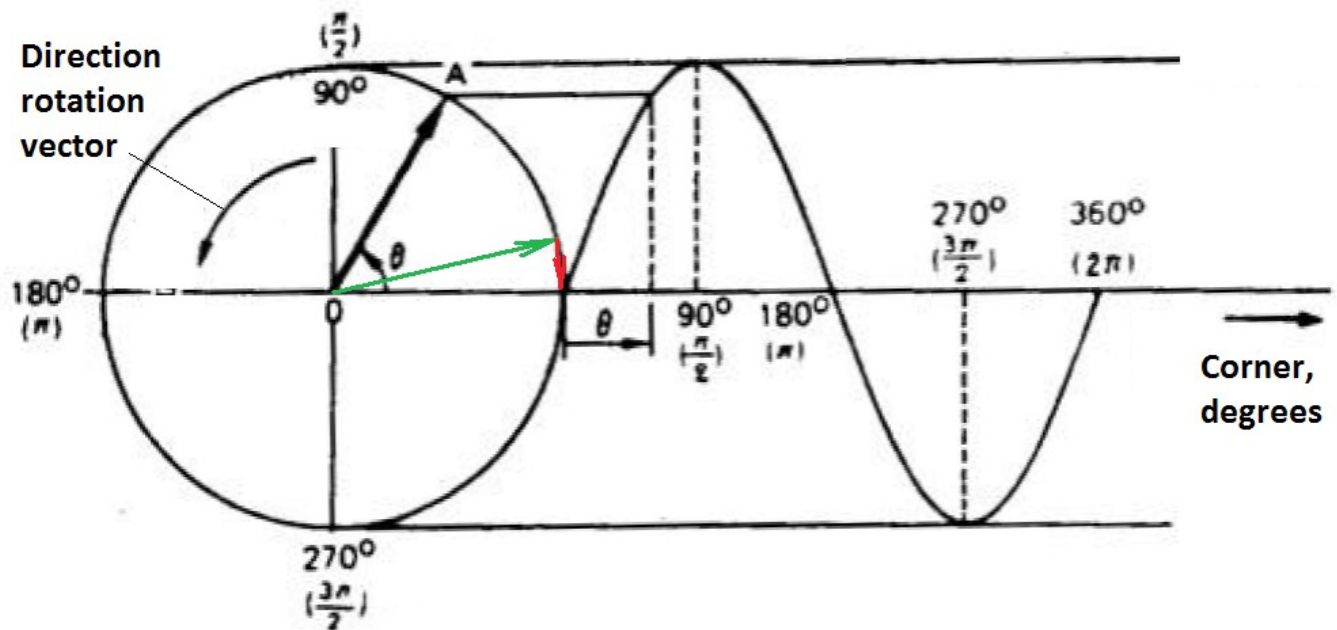


Fig. 16. Illustration of the mechanism of formation of a sinusoidal oscillation using a rotating vector.

The following illustrations show the relationship between vector error and propagation delay time. Datasheets for some high-speed op-amps list the propagation delay time as tPD (time Propagation Delay).

Figure 17 shows the ideal model of vector addition of sinusoids.

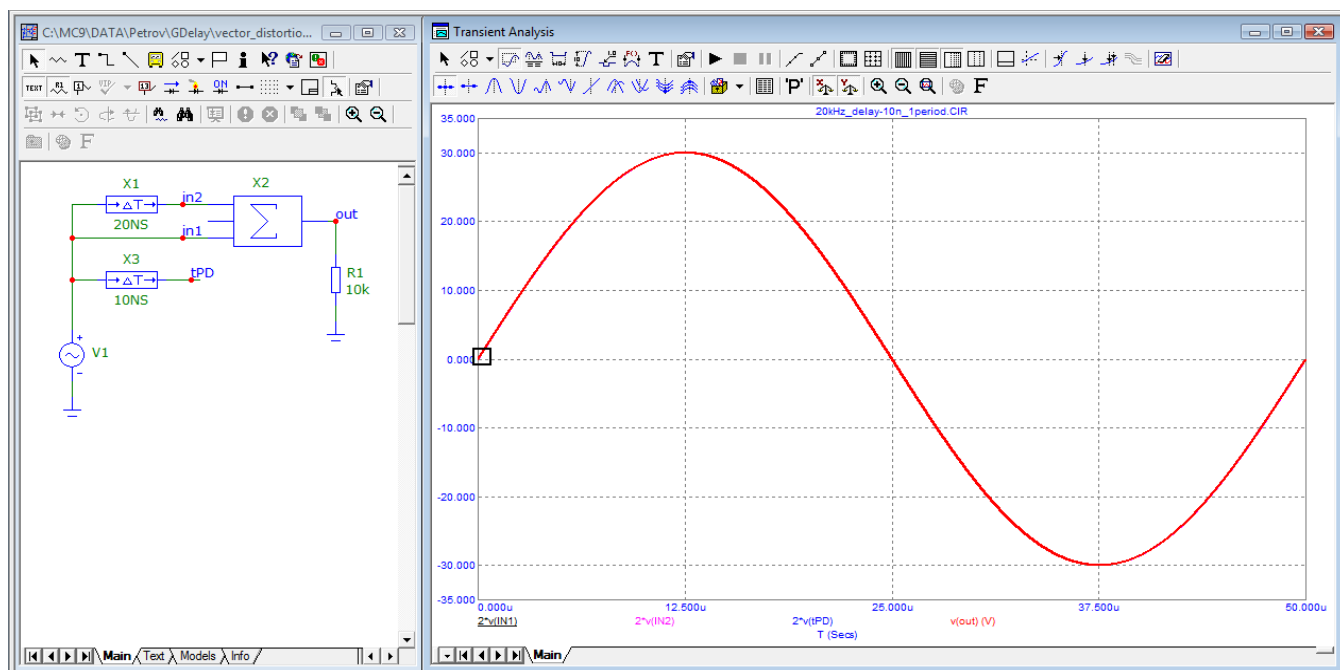


Fig. 17. Stress vector addition model

Figure 17 shows four sinusoids, but since they are delayed relative to each other for a short time, they visually merge. The oscillator signal with a frequency of 20 kHz has an amplitude of 15 V. When two signals are added, one of which is delayed for a short time, the amplitude of the output signal almost doubles and is equal to 30 V. In order for all four signals to be of the same amplitude, a multiplier of 2 is applied to the signals IN1, IN2 and tPD .

The initial section of the signal highlighted by the rectangle is shown in the following figure in a stretched form fig. 18

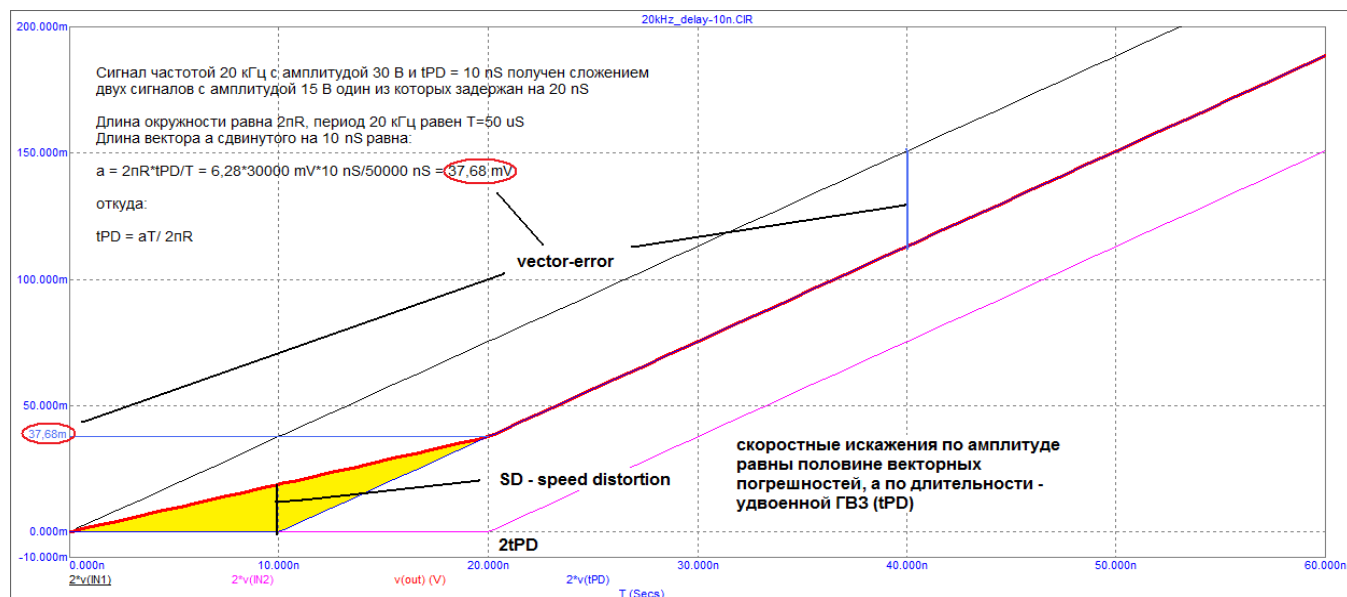


Fig. 18. The result of vector addition of 2 ideal sinusoids (initial section)

Figure 18 shows that when adding the original signal (black) and the signal (pink) delayed by 20 ns, an output signal (red) is formed with a delay of half that, i.e. 10 ns. Moreover, in the initial section, distortions inevitably occur over 20 ns, i.e. over $2 \cdot tPD$. This type of distortion appears only at the beginning of the period of addition or subtraction of two ideal signals. I took the responsibility to call this

type of distortion speed, since it is directly related to such a speed parameter as t_{PD} and occurs at the moments of change in speed dV / dt . Ideally, these distortions are equal in amplitude to half the vector error (from the projection of the yellow triangle onto the horizontal axis), and in time $2 \cdot t_{PD}$. The amplitude and duration of high-speed distortions (yellow triangle) determines their contribution to the overall level of distortions of complex signals, which include the audio signal. It is this kind of distortion that makes up the bulk of the distortion of the first period of Graham Maynard known as (FCD).

The figure also shows a calculation in accordance with [12] showing a direct dependence of vector errors and high-speed distortion on t_{PD} .

In order to make the most of this ideal model, let us measure the speed distortions with its help, Fig. 19

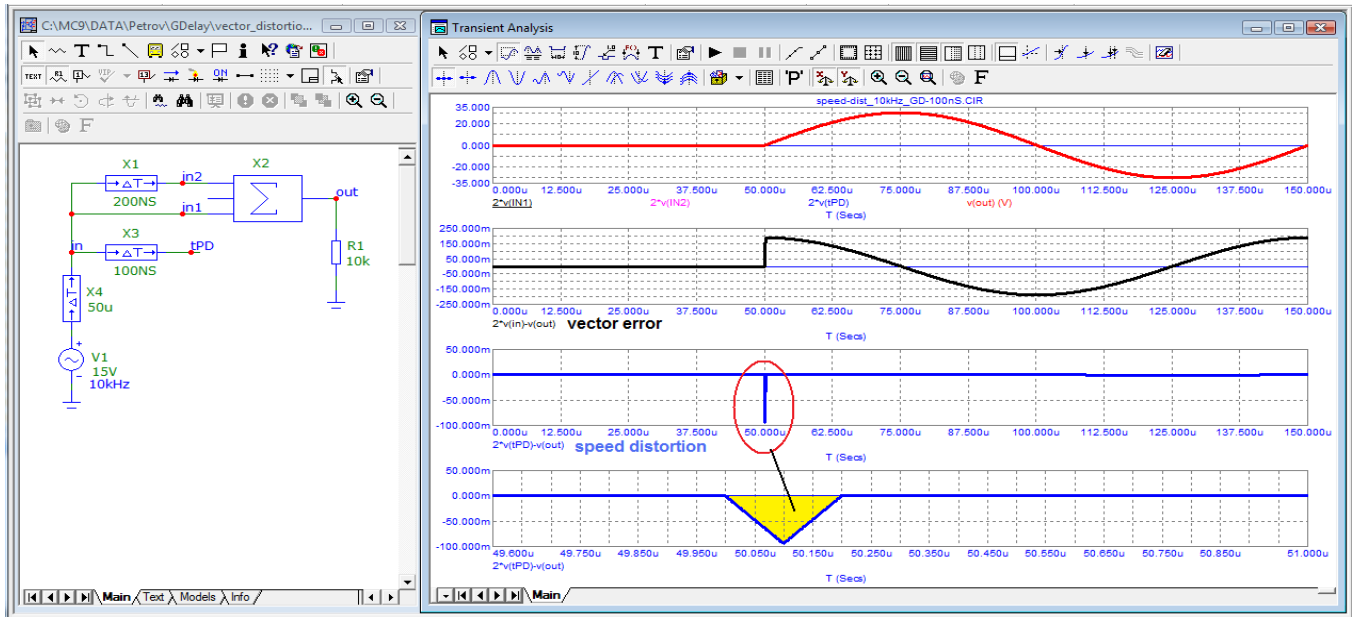


Fig. 19. The result of measuring the distortion resulting from the vector subtraction of two ideal phase-shifted voltages.

It can be seen from Figure 19 that when the output signal, delayed by 100 ns, is subtracted from the input signal, distortions occur with a negative sign with respect to the polarity of the first half-wave of the input signal and an amplitude equal to half the vector error and a total duration of 200 ns, i.e. $2t_{PD}$. These distortions have nothing to do with linear distortions (neither in phase nor in amplitude). **These are the distortions I refer to as high-speed ones.**

A sinusoidal voltage is formed by a vector of constant magnitude rotating at a constant angular velocity (see Fig. 16). With any change in both the amplitude of the vector and the angular velocity of its rotation, the voltage dV/dt also changes. At these moments, additional distortions associated with the signal propagation delay time (t_{PD}) occur.

Note. In formulas (6) and (7) there is an approximate sign, since strictly speaking the chord and the arc are not equal in length (Fig. 16). But for the highest frequency of the audio range, the period is 50 μs , and the signal propagation delay time rarely reaches 1 ... 2 μs . Therefore, even for such large values, the error is negligibly small.

For clarity of vector errors, let's assemble another model, fig. 20

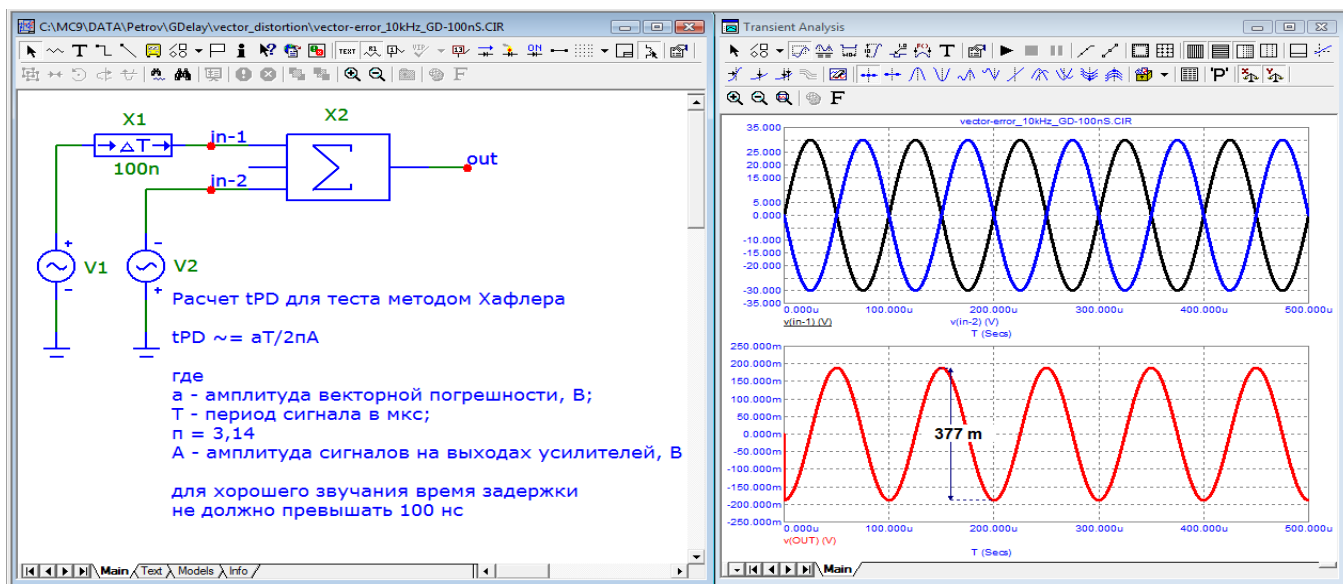


Fig. 20. The result of measuring the vector error of phase-shifted two voltages (for example, as in the SWDT test)

As a result of adding two signals in antiphase, one of which is delayed by 100 ns at the output, we obtained a vector difference in the form of a sinusoidal signal of the same frequency, the amplitude of which obeys formula (6):

$$a \sim 2\pi A * tPD/T = 6,28 * 60 * 0,1/100 = 0,377 B$$

In real amplifiers, the amplitude of speed distortion (not to be confused with speed errors or, what is the same, vector errors) can be somewhat lower than calculated - it depends on the value of tPD and its behavior both in the audio band and beyond it.

An example of the analysis of high-speed distortions of the Apex HD50 industrial amplifier model can be found in [16].

In order to isolate distortion products in more detail (in particular, switching distortions), it is necessary to use an ideal delay. But this can only be done in the simulator. In the gland, you can combine the Hafler test with the Sapozhkov test, Fig. 21.

A tunable delay line can be made with an acceptable error using a simple RC circuit whose time constant is tPD. The capacitor should be used high-voltage, for a voltage of at least 250 V, and the probe should be with a calibrated divider. To increase sensitivity when using a spectrum analyzer instead of an oscilloscope, you can use a wideband preamplifier. If measurements are made without a divider, then the calculated capacitance must be taken into account the capacitance of the connecting cable.

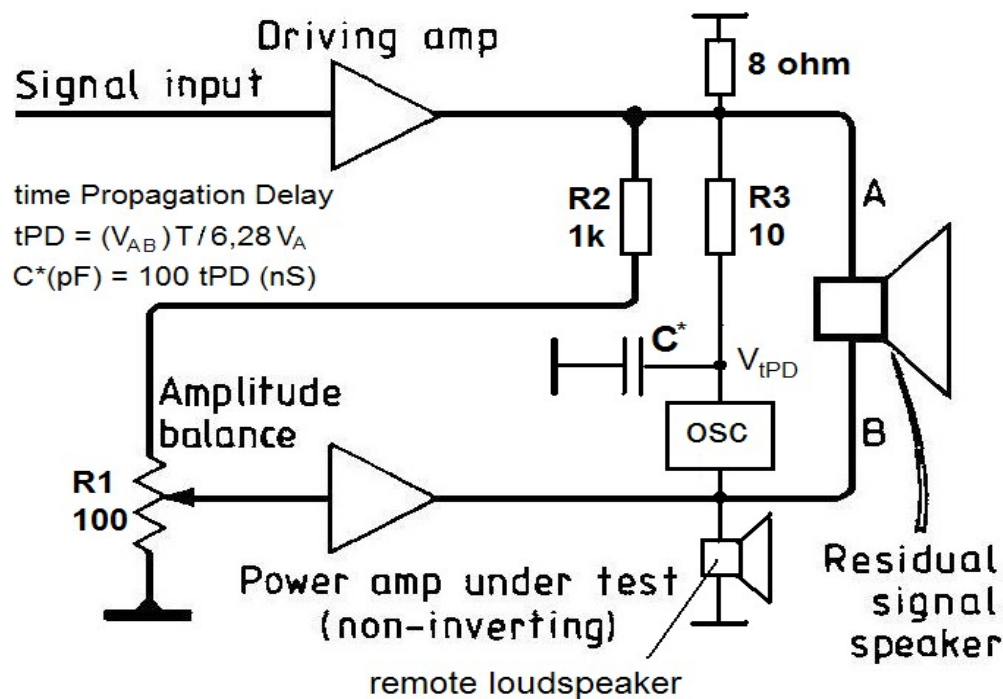


Fig.4. Hafler 'straight-wire' differential test.

Fig. 21. Block diagram of Hafler combined with Sapozhkov vector distortion analyzer

As a delay line, in some cases, you can use the Boucher chain of the upper amplifier. A typical Boucher circuit consists of a 10 ohm resistor and a 100 nF capacitor connected in series, the time constant of such a circuit is 1 μ s. As practice shows, the signal propagation delay time in real amplifiers varies widely: from tens of ns to 1.5 μ s or more. Therefore, it is quite possible to use the Boucher circuit of the upper amplifier to form a delay by the calculated value by replacing the capacitor and adjusting the resistor R3 to the optimal value, taking into account the input capacitance of the oscilloscope cable.

If the delay t_{PD} is not more than 100 ns, then the high-speed distortions are already negligible and can be neglected, they will be at the level of other types of distortions. If the delay is large (maybe 1.5 μ s or more), then such a delay will compensate for high-speed distortions, but other types of distortions (crossover, nonlinear) will still show, with the possible exception of distortions caused by amplitude-phase conversion, since phase and amplitude errors will be largely compensated by the delay line on the RC chain.

You can often hear from "theorists" that the manifestation of group delay in amplifiers is no different from the distortions introduced by a simple RC circuit. This can be said by someone who does not know that the RC circuit introduces a delay that smoothly drops to zero beyond the passband, while in amplifiers the behavior of the group delay (t_{PD}) is unstable and very diverse (often with significant spikes in the positive region and dips in the negative region).). It is the nature of group delay behavior that determines the duration and nature of transient processes and introduced distortions in the first periods (especially in the first half-cycle). Examples of group delay behaviors most frequently encountered in practice in UPT are shown in fig. 22

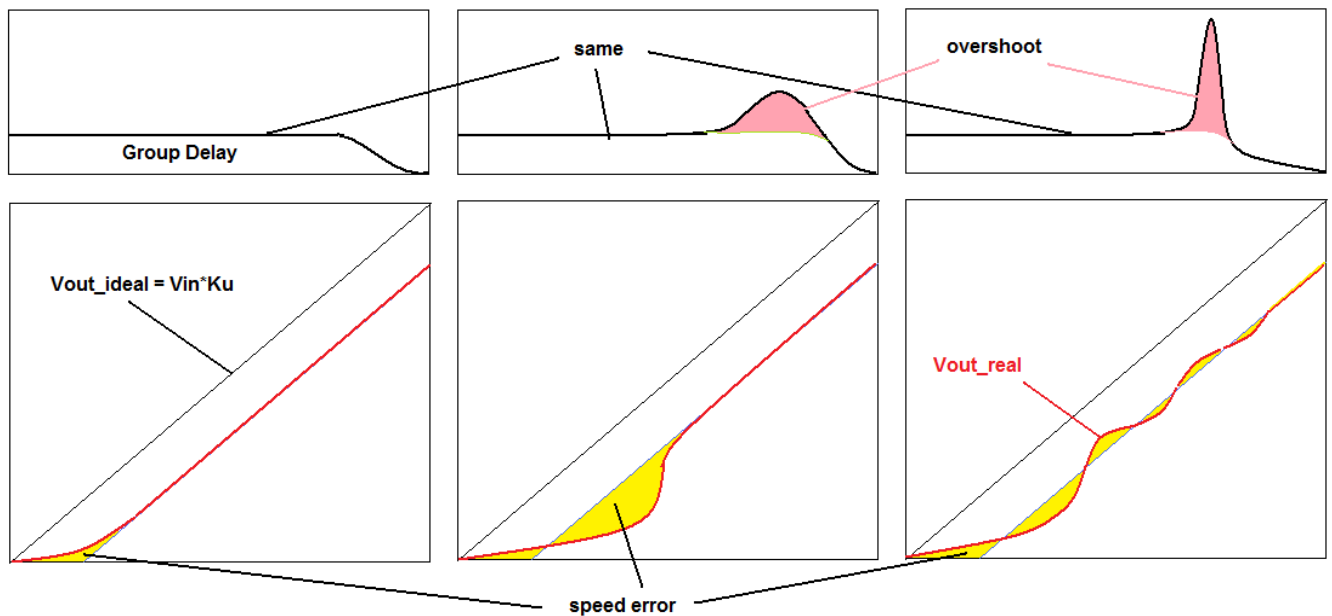


Fig. 22. Examples of Group Delay Behaviors Encountered in Practice in DC Amplifiers

In amplifiers with a closed input (with a coupling capacitor at the input and in an NFB divider), in amplifiers with voltage additives, and also in DC Amplifiers with a servo control system, the behavior of the group delay is much more complicated, often with a “fly away” to the negative region.

Figure 23 shows the dependence of high-speed distortion on group delay, and also what is the difference between vector errors and high-speed distortion of two amplifiers

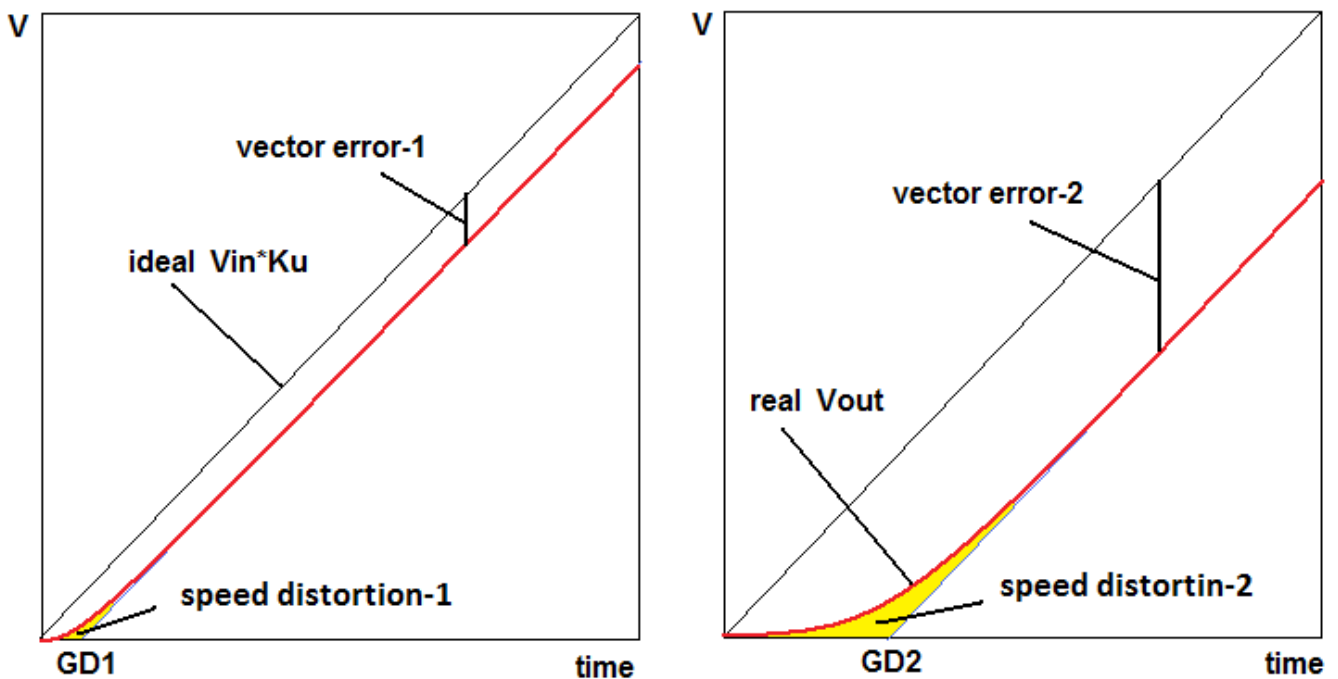


Fig. 23. Two examples of the dependence of speed distortion on group delay

Note. The most interesting thing is that even the beginning of the burst, even the beginning of the triangular signal after processing the 100 kHz low-pass filter; does not have a high slew rate (SR), on the contrary, the slope of the signal, and, accordingly, the SR is significantly lower than in other areas. Therefore, this kind of distortion has nothing to do with SID type distortion when the slope of the signal approaches the limiting SR of the amplifier. It is generally accepted that in Hi-Fi amplifiers SR should be at least 50 V / μ s. It is known that the maximum slew rate of real musical signals at the output of amplifiers rarely exceeds 0.5 V/ μ s, i.e., an order of magnitude lower. Therefore, the occurrence of SID-type distortion in amplifiers is unrealistic.

Nevertheless, it is in these areas (at the beginning and end of the burst) that signal losses with an amplitude of up to 1% or more occur. It is this type of distortion that is responsible for the loss of microlevel information that determines the "liveness" of the sound.

The Hafler test and the Hafler-Sapozhkov test are based on subtracting the scaled original (normalized input) signal from the output signal. The absence of any (audible) residue proves that there are no mysterious, unmeasurable ingredients. Both methods can be successfully used to evaluate the performance of amplifiers.

The advantages of the latter methods include the possibility of using any signals as test signals, including musical ones, as well as their extreme simplicity, which does not require expensive special measuring equipment, which is no less important.

When using a triangular signal during balancing, a differential voltage close to a meander is achieved. In this case, the meander amplitude will depend on the signal propagation delay tPD as follows:

$$a = 4A(tPD)/T \quad (9)$$

where

T - is the signal period, μ s;

A - is the signal amplitude at the amplifier outputs, V;

tPD – signal propagation delay time, μ s

When using a test signal in the form of a meander, the difference signal should look like a straight line on which bursts of up-down amplitude occur opposite the signal fronts and the duration of which depends on tPD.

Note. *To avoid overloading the amplifier input, test signals are passed through an RC filter with a cutoff frequency of 100 kHz (as in the DIM-100 test). The vast majority of amplifiers have a similar filter at the inputs with a cutoff frequency of 160 kHz (1 kOhm, 1 nF).*

To quantify the introduced high-speed distortions (measure their amplitude and duration), it is best to use an oscilloscope.

Today, the design of amplifiers does not occur without their preliminary simulation. Vector errors can be measured according to Dostal, and all other types of distortion (nonlinear, crossover and high-speed) - by a compensation method using an ideal delay line equal to the signal propagation delay time at the tested frequency. When making measurements, to improve accuracy, careful tuning should be done to determine both the gain and tPD as accurately as possible.

Methods for measuring dynamic nonlinear distortions [17].

One of the widely used methods for measuring dynamic non-linear distortion is the so-called "sine-rectangular signal" method. The measuring signal is a sequence of rectangular pulses with a repetition rate $f_1 = 3.18$ kHz, on which a harmonic signal with a frequency $f_2 = 15$ kHz is superimposed. The signal voltage amplitude ratio is $U_1 : U_2 = 4:1$, the signal slew rate is limited by the low-pass filter on RC links, the cutoff frequency of the filter is $f_{LP} = 30$ and 100 kHz (respectively DIM-30 and DIM-100). The measure of non-linear distortion is the ratio of the effective values of the intermodulation products to the amplitude of the harmonic signal.

Another method is based on the use of a measuring signal, which consists of two harmonic signals of the same amplitude, slightly different in frequency. Such frequency pairs can be, for example, $f_1 = 14$ and $f_2 = 15$ kHz or $f_1 = 19$ and $f_2 = 20$ kHz (CCIF-IM). The measure of distortion is the ratio of the effective values of the amplitudes of the intermodulation products $m_{f1} \pm n_{f2}$ to the amplitude of the main signals.

The third method - the method of sawtooth signals - is based on the use of a sawtooth signal with a pulse repetition rate $f_1 = 20$ kHz, which periodically (with a frequency $f_2 = f_1 / 256$) changes polarity. Distortions caused by a limited signal slew rate appear as periodic (with a frequency f_2) shifts in the average values of the DC voltage component. The measure of distortion is the ratio of the levels of this low-frequency voltage and the sawtooth signal.

It should be noted that not all testing methods are considered. For example, the review did not include Hiiragi's inverse intermodulation method (RIMD) [18], as well as Hirata's testing method [19]

As for the use of NFB to reduce distortion, there are several authoritative statements about this. For example, Cyrill Hammer said the following in an interview [20]:

“Perfect performance is just as important. This is especially true for amplifiers with a common NFB. The theoretical concept of negative feedback is very powerful, and the simplified mathematical equations that describe this concept always hold true. But they are only valid if the design takes into account the limitations of the concept. Delay from input to output must be zero! Obviously this is not possible in real life.

There are two ways to solve this problem:

1. You simply don't use any negative FB at all in your project (giving up the benefits of the concept).
2. You speed up the amplifier to a few nanoseconds of input to output delay (respectively 200 MHz bandwidth...), then the compensation errors are so small that they don't have any noticeable effect on the sound.

As soon as you decide to go the second way, a lot of new problems suddenly arise: thermal conditions, supply voltage stability, design suitable for high-frequency devices, the occurrence of noise, interference, etc.”

Here it is appropriate to recall the statement of Martin Colloms [21]:

“Engineering can explain a lot about the world of sound. However, when that fails to explain something, the real fun begins. Some aspects of perceived sound quality are not explained by established theory. There is growing suspicion that some of these aspects are a loss of natural timbre; boring, less expressive performance; increased auditory fatigue; and the lack of life and energy in the reproduced sound - may be a consequence of the application of negative feedback.”

Emeritus Professor Malcolm Hawksford paid much attention to the first path, proposing highly linear amplification stages with distortion compensation, including those with positive feedback (PFB) for the output current [22].

Manufacturers producing two types of amplifiers are constantly competing with each other. And although the first amplifiers formally have a higher level of distortion measured on a sinusoidal signal in a steady state, in subjective tests they most often turn out to be the winners.

For example, Jeff Rowland believed that "OOS amplifiers introduce distortion in the time domain" therefore he developed the "model 7" non-OS amplifier in three modifications that differ in both voltage amplifiers (VU) and output stages (VK) - technical solutions are not disclosed. A characteristic feature of these models is the constancy of the output impedance over the entire sound range, an extremely high load capacity (up to 150 A-peak) and an adequate response (in phase) to external influences in the form of back-EMF acoustics. And although they had slight differences in sound, they nevertheless received high marks from professionals and lovers of high-quality sound.

Suffice it to say that individual components of the company cannot be purchased even on the secondary market - amplifiers such as "Coherence", Model 7 and Model 9 have become rarities and a living legend among music lovers.

The second path is used by the legendary designer and consultant for several audio companies, John Curl. In doing so, he adheres to the following rules:

- output voltage slew rate not less than 100 V/μs;
- the frequency of the first pole is as high as possible in the sound range (preferably not lower than 100 ... 200 kHz);
- minimal feedback, and its absence is better;
- operation of the output stage in class A up to 10 W or more (mitigates switching distortion);
- as low as possible the level of high-order odd harmonics starting from the 5th and higher;
- if possible, avoid the use of an output choke to ensure stable operation, since a choke with an inductance of even a fraction of μH introduces a significant delay and, in combination with a reactive load, can create oscillatory transients;
- quality power for each cascade.

He also shared an experiment he demonstrated to his collaborators about 40 years ago. Otala's Electrocompaniet was used as an amplifier, and a Pioneer ribbon tweeter with a bandwidth of more than 45 kHz was used as a tweeter. As a test signal, he used a rectangular signal (meander) with a frequency of 5 kHz, processed by switchable first-order low-pass filters of 35 and 100 kHz. In this case, the rise time of the fronts changed from 10 μs to 3.5 μs. All employees heard the difference in sound. From this simple

experiment, he deduced that human hearing is more sensitive to the rate of change of a signal than to the actual frequency response.

Более подробно с его идеологией можно ознакомиться в [14], а также в ряде высказываний на форумах, в частности [23].

More details about his ideology can be found in [14], as well as in a number of statements on forums, in particular [23].

For example, in an interview [20] Nelson Pass said the following: "The last ten years have shown once again that high quality amplifiers with per millionth distortion and other excellent performance are not very popular. It's like pure distilled water - it has no taste and most people don't want to drink it. »

Therefore, it is not surprising that many well-known manufacturers such as: Ayre, Akai, Denon, darTZeel, Denset Beat, Krell, Lamm, Pass Laboratories, Pioneer, Rowland, Sony, Tandberg, Threshold and others produced models without a common NFB. And companies like NAD began to raise the frequency of the first pole above the audio range, while limiting the loop gain.

In one of the interviews, V. Shushurin (Vladimir Lamm) gave the following example: "We have three amplifiers: one has 1-2% distortion, the other has 0.1% distortion, the third has thousandths of a percent. We listen to all this through a speaker system that has 5% distortion. In theory, we should not hear the difference - only the handwriting of the loudspeaker. But we can hear perfectly well that the amplifiers sound different. Paradox, right? The answer to this question is not so easy to find. It took me several years to figure out how to answer it." True, he did not share the answer ...

According to the Rakovsky scale [2], THD is at the lowest level of correlation with sound quality. Much more important is the proportion of higher odd harmonics starting from the 5th and higher (especially the 7th and 9th).

John Curl, referring to the 1941 German "Handbook of Radiotron Engineering" gives the following harmonic weighting factors [14]: $N^2 / 4$ (N to the power of 2 divided by 4) for each harmonic. In this case, the second harmonic has a weighting factor of 1, and for example, the 7th harmonic is already 12.5 (22 dB higher).

Moreover, THD does not take into account such a high-speed parameter as the signal propagation delay (time Propagation Delay) and its behavior far beyond the audio range. But it is the group delay that has the greatest impact on the sound quality. The IRI measures distortion in the steady state (i.e., at the end of transients) and does not take into account distortion in the time domain (velocity and transient distortion).

As for the effect of harmonic components on the timbre of sound, this was described more than 80 years ago [24]:

"The second harmonic adds clarity and brightness, but nothing more, since the general principle is that adding an octave cannot make any difference in timbre or distinctive musical quality. When the second harmonic has the same strength as the first, it has almost the same effect as adding an octave connector on an organ or harmonium, or playing in octaves instead of individual notes on a piano.

The third harmonic again adds some brightness due to its high tone, but it also makes a difference in timbre, thickening the tone and adding to it a breathy or nasal character that we can recognize as one of the main ingredients of the clarinet tone.

The fourth harmonic, being two octaves higher than the fundamental, adds even more brightness and perhaps even shrillness, but no more than that, for the reason already explained.

The fifth harmonic, in addition to adding even more brilliance, gives the tone a richness, somewhat similar to a horn, and the sixth adds a subtle piercing nasal quality.

All these six harmonics form parts of a common root note chord and therefore agree with that note and with each other.

However, the seventh harmonic introduces an element of dissonance. The same is true for the ninth, eleventh, thirteenth and all higher odd harmonics; they also add dissonance as pitch harshness, and thus bring roughness or harshness to the compound sound. The resulting tone quality is often described as "metallic"

Back in the 1950s, Theo Williamson wrote that for a high-quality amplification of an audio signal,

it is sufficient that K_r be no more than 0.1% at maximum power. In this case, the harmonic content is virtually undetectable in the most sophisticated listening tests (apparently he was referring to short-range tube amplifiers, this is not the case with modern transistor amplifiers). He also noted that phase shifts between the harmonic components of a complex signal in dynamics, especially on attacks of sounds, have a significant impact on the naturalness of the sound.

Low first pole amplifiers are known to have a loop gain phase of 90 deg. in almost the entire audio range. Here is what S. Ageev writes about this [25].

“When the phase angle of the loop amplification is close to ± 90 or ± 270 degrees, the amplitude non-linearities of the original amplifier are almost completely converted into phase ones (i.e., into parasitic phase modulation, albeit weakened by $|bK|$ times). In this case, parasitic amplitude modulation practically disappears, and the results of intermodulation distortion measurements can be 20 ... 30 dB more optimistic than the spectrum analyzer (and hearing in the case of a power amplifier) actually shows. Unfortunately, that's the way it is with most ops and many AMPs.”

There have been attempts to artificially widen the region of constant loop gain by using a two-pole correction. But this only gives an improvement in the measured parameters (THD), the sound quality only deteriorates due to changes in the group delay characteristic.

Regarding the operation of NFB in audio frequency amplifiers, Charlie Hansen writes the following [26]:

“In terms of using feedback in audio circuits, we already know it doesn't work as advertised. The idea is that the output of the circuit is compared to the input. Any differences (noise or distortion of any type) are subtracted from the incoming signal. But if this process worked correctly, then all amplifiers would sound the same, since they would all perfectly reproduce the input signal.

Unfortunately, the reality is somewhat more confused than the theory. In fact, every amp sounds different. Applying negative feedback does not help this in any way.

The differences between the sounds of amplifiers with feedback are no less than those of amplifiers without feedback. In fact, in general, our experience is that zero feedback amplifiers tend to sound more similar to each other than typical feedback amplifiers.

A much better approach is to simply design a circuit that is inherently linear, to the point where the distortion contribution from the amplifier is significantly (eg 10 times) lower than that from the loudspeakers. When this condition is met, there is no reason to apply feedback. Anyone familiar with the sound of zero-feedback amplifiers will find it hard to go back to traditional designs.”

Distortion during transients.

The harmonic signal is described by the following formula:

$$A \cos(\omega_0 t + \varphi_0)$$

where

A - is the signal amplitude,

$\omega_0 = 2\pi f_0$ – circular frequency,

φ_0 - is the initial phase.

In this case, A is a constant value, and the signal spectrum consists of a single component with a frequency ω_0 - a monochromatic spectrum.

During transients, the signal becomes quasi-harmonic [27]:

$$s(t) = A(t) \cos[\omega_0 t + \varphi(t)]$$

In contrast to the steady state, the amplitude of the signal A and the initial phase φ depend on time in the section of the transient process. In this case, the signal amplitude A is not constant (as with a harmonic signal), and the signal spectrum becomes complex (with additional harmonic components) depending on the behavior of the function A(t). The first period of the burst is subject to the greatest distortions (including high-speed ones due to the finite signal transit time). In some cases, such distortions are visible to the naked eye, especially on a reactive load. As the transient processes decay (theoretically, their duration is infinity), the distortions decrease. However, already from the second period, they are practically invisible to the naked eye, and high-speed distortions are completely absent.

With regard to audio frequency amplifiers, DC amplifiers are in a more advantageous position. The reason for additional transients (except for coupling capacitors at the input and in the FB

circuit, as well as servo control systems) are often inductances included at the output of the amplifier designed to ensure stable operation. These inductances form a series oscillatory circuit with a reactive load in the form of a capacitance reaching in some cases 8 μF . Therefore, it is advisable, if possible, to develop amplifiers that do not need to stabilize their operation with the help of inductance. In some cases, an inductance of 0.1 ... 0.2 μH is sufficient for stabilization. Suffice it to say that a wire with a diameter of 1 mm and a length of 10 ... 20 cm coming from the board to the output connectors has such an inductance.

A few words about the signal propagation time (time Propagation Delay). In DC amplifiers, the group delay (GDT) of the signal has a horizontal section from infra-low frequencies and covers the entire audio range. It is very important that this section is not less than 300 kHz. Above this frequency, the group delay should have a smooth decay to zero. A slight increase in group delay outside this range is acceptable. Significant group delay rises contribute to the growth of high-speed distortions (distortions that occur when the frequency or amplitude of the signal changes). The easiest way to measure them is by the compensation method using test signals with a frequency of 10 kHz (triangular or sinusoidal bursts).

On the horizontal sections of the group delay, its value coincides in magnitude with the value of tPD. The value of vector errors (a) within the sound range for a sinusoidal signal is calculated using a simplified formula (6) [12].

The criterion for performing the SWDT test (D. Hafler) is a vector error of no more than -60 dB at a frequency of 10 kHz, and no more than -70 dB at the middle frequencies of the audio range [13]. This requirement corresponds to $tPD = T/1000 \cdot 2\pi = 100000/6280 = 16 \text{ ns}$ for a frequency of 10 kHz, and 8 ns for a frequency of 20 kHz.

However, as practice shows, in some cases it is possible to obtain quite good results in sound quality with tPD no more than 100 ... 120 ns ((for example, MIMESIS 9.2 - depends on the group delay behavior above 300 kHz) and there is no need to use an inductance at the amplifier output.

As for the speed parameters of modern amplifiers, in particular the slew rate, for example, the top NAD M3 amplifier has a slew rate (SR) of 1000 V/ μs ; HK Citation XX - 500 V / μs ; Denon POA 2400 - 500 V/ μs ; HK PA-2400 - 280 V / μs ; KR-8050 - 200 V/ μs ; KR-770 - 180 V/ μs ; HK Citation 22 - 160V/ μs ; Sansui G-22000 - 175V/ μs etc. - the list goes on and on...

Here is some more reasoning about distortion and the operation of NFB in amplifiers.

Harmonic distortion [28].

“Manufacturers indicate only the total Kg (root sum of squares of individual harmonics), while nothing is said about the spectral composition of these distortions. That is, an amplifier with a total Kg = 0.01% can subjectively be both very mediocre if the high harmonics are not sufficiently suppressed, but it could also be quite decent (if the lower harmonics make the main contribution).

It is wrong to talk about distortions, considering only harmonic ones, regardless of intermodulation ones.

The fact is that the same nonlinearities in the amplifying path that generate harmonics, with absolute inevitability, also generate intermodulation. And this is not a subject for discussion, this is a mathematically proven fact. In fact, harmonic distortion is just a special case of intermodulation, when one of the test frequencies is missing. The intermodulations of high-frequency components also fall into the middle frequencies, in the zone of the greatest hearing sensitivity, and are not masked by high-frequency components! The hearing threshold at mid frequencies is around 0 dB and it is important that the intermodulations are kept below this threshold. First-order intermodulation can be equal to harmonics in amplitude, hence the requirement: the level of harmonic distortion at high frequencies of the entire path (especially difficult to achieve in the PA) should not exceed the threshold of audibility at medium frequencies.

Thus, for a sound pressure of, for example, 96 dB, the level of harmonic distortion at HF should not be more than -96 dB (0.0016%).

It is highly desirable (especially for PA) to have spectrograms for different signal levels. Most importantly, such spectrograms make it possible to judge the nature of the decrease in harmonics. It has been proven that human hearing is insensitive to the second and third harmonics, but has an amazing sensitivity (up to 0.001%!!! - 100 dB) to high harmonics, starting from the fifth.

Incorrect, "harmful" work of NFB.

The output signal turns out to be delayed in time (strictly speaking, this is no longer the same signal that is currently acting at the input). It is no longer clear what we are subtracting from. This difference (error signal) is unpredictable, and in this case the amplifier frantically tries to follow the input action that is difficult for it, it is on the verge of stability.

The lower the delay, the more benefit and less harm from NFB.

Matti Ojala argued in his 1970 work that "the slew rate must exceed the rate corresponding to the bandwidth of the amplified frequencies by a factor depending on the FB depth and equal to at least 50, otherwise impulse intermodulation distortion, T.I.M.

Calculated according to this statement, the required slew rate for 20 kHz and 100 W into 8 ohms is 248.8 V/ μ s, and the bandwidth is 1 MHz! How many, even ultra-modern amplifiers, can boast of such a speed?! Very few, but they exist!

The speed margin, together with good linearity and a large gain margin, creates a very favorable forecast for the smallness of various signal distortions, including high-order intermodulation and harmonic distortions. The latter are most noticeable to our ears and are responsible for the "transistor" sound. "

I hope that the presented material will allow us to take a fresh look at the long-known information both on the requirements for high-quality sound amplification amplifiers and on their testing methods to obtain the maximum correlation of measurement results with sound quality.

Literature:

1. K. Kupfmüller. Fachberichte des Jahresammlung, Vers, d.V.d.E, 1926, S.87
2. В.В.Раковский, Измерения в аппаратуре записи звука кинофильмов, «Искусство», Москва, 1962
3. Lynn Olson, Matt Kamna, The Sound of the Machine. The Hidden Harmonics behind THD, «Glass Audio» 4/97
4. В. А. Симон и др. Режекторный фильтр с настраиваемыми параметрами, Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В. И. Ульянова (Ленина), УДК 621.396.6; 621.375.13
5. В.М. Вольф, Способ измерения нелинейных искажений, патент SU90185
6. И.Алдошина, 126 конгресс AES в Мюнхене - научные результаты, Звукорежиссер 2009-07
7. A.Darre, Methoden zur Messung nichtlinearer Verzerrungen im Tonfrequenz Gebiet Frequenz, 1955 №3, 4
8. А.Сырицо, Измерение нелинейных искажений на шумовом сигнале, Радио 1999, №4
9. М.А. Сапожков, Корреляционный метод измерения нелинейных искажений передачи, Акустический журнал, Том 2, вып. 3, 1956
10. Н.Л. Безладнов, Об усовершенствовании методов измерения нелинейных искажений, «Электросвязь», 1957, № 2.
11. P.J. Baxandall, Audible amplifier distortion is not a mystery, Wireless World, november 1977
12. И.Достал, Операционные усилители, Москва, «Мир», 1982
13. D.A. Hafler, Listening Test for Amplifier Distortion, Hi-Fi News and Review, November 1986, pp.25-29
14. John Curl, Condemnation without Examination is Prejudice, or Words of Wisdom, 2006
15. J. Gordon Holt, The Carver Challenge, First Published: Oct 11, 1985 (<https://www.stereophile.com/content/carver-challenge>)
16. А.Петров, Скоростные искажения, Радиоаматор 2020 №5-6 с.31
17. П.Шкритек, Справочное руководство по звуковой схемотехнике, Москва «Мир» 1991
18. J.Hiraga, The Many Faces of Distortion, Glass Audio, may 2005
19. Y.Hirata, Nonlinear Distortion Measurement Using Composite Pulse Waveform, J. Audio Eng. Soc., Vol. 29, No. 4, 1981 April
20. The Absolute Sound_ May/June 2012 (<https://www.moremusic.nl/reviews/passlabs/XP-30-TAS.pdf>)
21. Martin Colloms, A Future Without Feedback? January, 1998, vol.21 No.1, p.87

22. Malcolm Hawksford, A Compilation of Technical Papers on Audio Amplifier Systems
23. <https://www.diyaudio.com/forums/solid-state/94676-bob-cordell-interview-negative-feedback-post1272373.html>
24. Sir James Jeans, Science and music, New York: At the UNIVERSITY PRESS, 1937
25. С.Агеев, Вопросы проектирования усилителей с общей ООС. — Радио, 2003, № 4, с. 16—19.
26. The AX-5 Story from Ayre Acoustics March 2014
27. Л.М.Финк, Сигналы, помехи, ошибки... , М, Радио и связь, 1984
28. <http://www.ecosound.pro/Distortions.html>
29. <https://www.theabsolutesound.com/articles/charlie-hansen-1956-2017/>
30. <https://www.youtube.com/watch?v=UxTmGzg-1rg>
31. С.Лозицкий, УМЗЧ с токовой обратной связью, Схемотехника, 2003, №2, с.12
32. И.Рогов, УМЗЧ с полевыми транзисторами в выходном каскаде, Радио, 2016, №10, с.8

material prepared
Petrov Alexander