

ПРЕОБРАЗОВАТЕЛЬ АНАЛОГ-ЦИФРА ДЛЯ
ВВОДА РЕЧЕВЫХ СИГНАЛОВ В ЭВМ

Г.Я. Волошин

Для проведения исследований звуков речи с помощью ЭВМ [1] необходимо представить исходный речевой сигнал в цифровой форме. Преобразователь Аналог-цифра, описываемый в данной работе, предназначен для выполнения этой операции. Все рисунки (I - I2) даны в приложении.

Выбор технических характеристик

а) Ч а с т о т а к в а н т о в а н и я. Как известно из теоремы Котельникова [2], частота квантования сигналов с ограниченным спектром должна быть не меньше удвоенной верхней частоты спектра входного сигнала

$$f_{кв} \geq 2F_{max}$$

Практически под F_{max} понимают или верхнюю частоту так называемой "эффективной" полосы [3], или частоту среза входного фильтра.

После ввода в ЭВМ сигнал подвергается различным преобразованиям, в том числе спектральному анализу. При этом используются различного рода весовые функции времени [4]. Ограничение сигнала во времени с помощью весовой функции эквивалентно расширению спектра этого сигнала. Частоту квантования сигналов, преобразованных с помощью единичной весовой функции вре-

мени, можно определить по формуле:

$$f_{кв} = 2F_{max} + \frac{0,1}{\tau_0 |\varepsilon_0|}$$

где τ_0 - минимальная протяжённость функции корреляции исходного сигнала;

$|\varepsilon_0|$ - абсолютная величина допустимой погрешности [5].

Практически для речевых сигналов $F_{max} \leq 10$ кГц. Выбор τ_0 меньше 5 мсек ведёт к появлению значительных погрешностей расчёта спектральной функции (см. [5]). Приняв $|\varepsilon_0| \leq 0,01$, $F_{max} = 10 \cdot 10^3$ гц и $\tau_0 = 5 \cdot 10^{-3}$, получим

$$f_{кв} \geq 22 \text{ кГц.}$$

б) Ч и с л о р а з р я д о в. Для анализа речевых сигналов нужно иметь устройства преобразования с динамическим диапазоном 45-48 дБ [6], что соответствует 9-разрядному преобразователю аналог-цифра.

Для удобства согласования преобразователя с ЭВМ СО АН СССР пришлось ограничиться 8-м разрядами (4I, 9 дБ), которыми, видимо, можно удовлетвориться, если учесть, что системы телефонной связи с динамическим диапазоном 40 дБ обеспечивают хорошую разборчивость и натуральность звучания.

Итак, технические требования к преобразователю сводятся к следующим:

- а) частота квантования должна быть не меньше 22 кГц;
- б) число разрядов - 8.

Блок-схема преобразователя

Преобразователь работает по так называемому компенсационному принципу (рис. 1, 2). В исходном состоянии все триггеры T_1-T_8 находятся в положении 0 и преобразователь цифра-аналог даёт на выходе напряжение 0в. В момент времени t , приходит запускающий импульс (ЗИ), приводящий в действие одновибратор ($ОВ_1$). Передним фронтом импульса, снимаемого с $ОВ_1$ через дифференцирующий каскад ДК и формирователь Ф, триггер T_1 устанавливается в положение 1. Одновременно весь импульс с $ОВ_1$ подаётся на схему совпадений $\bigvee_{n=1}^N$. Преобразователь цифра-аналог вырабатывает напряжение $\frac{U_{max}}{2}$, которое сравнивается с входным $U_{вх}$. Сравнительное устройство выдаёт импульс для сброса триггера только при $U_{вх} < U_{y-a}$ где U_{y-a} - напряже-

ние, вырабатываемое преобразователем цифра-аналог (в рассматриваемом случае импульс сброса не будет и I-й триггер остаётся в положении "1"). Задним фронтом импульса, снимаемого с OB_1 , запускается OB_2 , и цикл повторяется во втором разряде.

Так, последовательно опрашиваются 8 разрядов, и в регистре T_1-T_8 устанавливается двоичный код входного напряжения. Затем производится съём показания регистра (t_9) и сброс триггеров в исходное состояние (t_{10}).

Преобразователь цифра-аналог (рис. 3)

Восемь генераторов тока нагружены цепной линией. Нетрудно показать, что сопротивление нагрузки любого генератора тока равно $2/3 R$. При замыкании поочерёдно ключей K_1, K_2, K_3, \dots напряжение на выходе цепной линии будет пропорционально величинам $1, \frac{1}{2}, \frac{1}{4}, \dots$

При одновременном замыкании нескольких ключей выдаётся соответствующее суммарное напряжение, если выполняется условие $R_i \gg R$, где R_i - внутреннее сопротивление генератора тока. Ключи K_1-K_8 управляются триггерами.

Принципиальные схемы (рис. 4)

Генератор тока. При выдаче с триггера или с тумблерного инвертора напряжения -10 в стабилитрон Д808 пробивается, напряжение в точке Б становится равным $-1,5$ в, диод D_1 запирается и генератор тока отключается от цепной линии.

При подаче с триггера напряжения -2 в диод D_2 запирается, генератор подключается к цепной линии. С помощью сопротивления R_2 компенсируется разброс параметров элементов цепной линии и генераторов тока. Для термостабильности U_{y-a} сопротивления в цепной линии и в эмиттерах триодов необходимо выбирать с одинаковыми ТКС. В этом случае термовариации тока триода сопровождаются равным и обратным по знаку изменением сопротивлений цепной линии, и напряжение на выходе не изменяется.

На рис. 5 и 6 представлены схемы одновибратора и триггера, разработанные в Институте математики СО АН СССР.

Триггер обеспечивает в счётном режиме быстродействие до 2 мкс. Управление ключом генератора тока осуществляется через дополнительный эмиттерный повторитель.

Одновибратор выдаёт импульс длительностью $2,0 \mu\text{сек}$. Отрицательный импульс подаётся на схему I_n , положительный - на дифференцирующие каскады (обычные RC - цепочки).

Схема I_n изображена на рис. 7.

Так как максимальное быстродействие триггера составляет 2 мкс, то между моментами установки триггера в положение 1 и сброса должен быть интервал времени не менее $0,5 \mu\text{сек}$. Этот интервал регулируется подбором ёмкости C^* в схеме I_n .

На рис. 8 изображена схема формирователя запуска (сброса) триггера. Ячейки I съёма кода - обычные диодные сборки, после которых стоят формирователи.

Выход синхронизирующих импульсов (СИ) осуществляется через усилитель и эмиттерный повторитель.

Сравнивающее устройство (рис. 9). На входе сравнивающего устройства стоит дифференциальный каскад, собранный на триодах T_1, T_2, T_3 . Положительное входное напряжение запирает триод T_2 по эмиттерной цепи. На базу T_2 подаётся U_{y-a} . Если $U_{y-a} > U_{bx}$, то T_2 отпирается, на его коллекторе появляется отрицательный импульс, который затем усиливается и формируется. Для триодов, применяемых в дифференциальном каскаде, обратное пробивное напряжение на базе равно 3 в. Поэтому, во избежание пробоя триода T_2 , диапазон входных напряжений выбран в пределах $0-3,06$ в. Чувствительность сравнивающего устройства не хуже 10 мв, при этом внутренние помехи преобразователя лежат ниже порога чувствительности.

Все связи в сравнивающем устройстве потенциальные, благодаря чему допускается любая скважность импульсов сравнивающего устройства (лишь бы эти импульсы не сливались друг с другом). Эмиттерный повторитель, собранный на триоде T_4 , диоды D_1, D_2, D_3 служат для стабилизации уровней импульсов.

Преобразователь осуществляет одно измерение входного напряжения за $16 \mu\text{сек}$, режим работы - "ждущий", т.е. один цикл преобразования осуществляется после прихода ЗИ.

В процессе настройки достигнута частота отсчётов 30 кГц. При незначительных изменениях схемы (нужно входы ячеек I_n со стороны сравнивающего устройства сделать потенциальными) частота отсчётов можно довести до 50 кГц.

Об искажениях, вносимых преобразователем

Оценим величины искажений, возникающих из-за того, что входной сигнал не запоминается на время одного цикла преобразования.

I. Амплитудные искажения. Для того чтобы не возникали амплитудные искажения в рамках заданной чувствительности преобразователя, необходимо обеспечить измерение любого значения входного напряжения в диапазоне 0-3,06 в.

Худшим случаем является измерение максимального значения $U_{вх}$.

На рис. 10 кривая I соответствует преобразуемому напряжению, II - эталонному, вырабатываемому преобразователем цифра-аналог.

Для измерения $U_{вх max}$ необходимо, чтобы за время $\Delta t = t_n^* - t_n$ кривая II достигла кривой I в точке a . Это выполнимо при условии, если кривая II целиком находится внутри кривой I и касается её только в точке a . Будем исходить из того, что на преобразователь подаётся напряжение с выхода стационарной линейной системы. Тогда анализ её амплитудно-частотной характеристики можно производить, подавая на вход синусоидальные сигналы разных частот и замеряя соответствующие выходные напряжения.

Пусть $U_{вх}$ линейной системы - синусоидальное напряжение. Необходимо выполнить следующее условие:

$$\Delta U_a \leq \Delta U_{min}, \quad (3)$$

где ΔU_a - изменение преобразуемого напряжения в окрестностях точки a за время $\Delta \tau$;

ΔU_{min} - величина ступеньки, соответствующей младшему разряду преобразователя цифра-аналог,

$\Delta \tau$ - временной интервал между ступеньками кривой II.

$$\Delta U_a = A \left[\sin \frac{\pi}{2} - \sin \left(\frac{\pi}{2} - \frac{\Delta \tau}{T} 2\pi \right) \right] = A \left(1 - \cos 2\pi \frac{\Delta \tau}{T} \right), \quad (4)$$

здесь A - амплитуда преобразуемого сигнала,
 T - его период.

От неравенства (3) перейдём к равенству:

$$\Delta U_a = \Delta U_{min} = A \left(1 - \cos 2\pi \frac{\Delta \tau}{T} \right), \quad (5)$$

отсюда

$$A = \frac{\Delta U_{min}}{1 - \cos 2\pi \frac{\Delta \tau}{T}}. \quad (6)$$

Так как $A_{max} = 127,5 \Delta U_{min}$, то

$$\frac{A}{A_{max}}(f) = \frac{1}{255(1 - \cos 2\pi f \Delta \tau)}, \quad (7)$$

где $\Delta \tau = 2,0 \cdot 10^{-6}$ сек. График функции $\frac{A}{A_{max}}(f)$ представлен на рис. II (кривая I). Так как A не должна быть больше A_{max} , то кривую I заменим кривой II, которую нужно понимать следующим образом: спектр сигнала, подаваемого на вход преобразователя, должен полностью уместиться внутри заштрихованной области. В противном случае будут иметь место амплитудные искажения. Или иначе: во избежание амплитудных искажений спектр

$$S_y(\omega) = |\Phi(j\omega)|^2 S_x(\omega), \quad (8)$$

где $S_y(\omega)$ - спектр сигнала, поступающего на вход преобразователя,

$\Phi(j\omega)$ - частотная характеристика стационарной линейной системы,

$S_x(\omega)$ - спектр сигнала, снимаемого с датчика,

может быть линейным до частоты 8 кГц, а затем должен убывать по закону, определяемому соотношением (7).

Зависимость $\frac{A}{A_{max}}(f)$ (рис. II, II) можно назвать предельным спектром преобразуемого сигнала, выход за пределы которого приведет к амплитудным искажениям.

2. Искажение формы преобразуемого сигнала. Допустим, что мы условились считать момент измерения момент запуска преобразователя, т.е. t_n (рис. 10). Тогда мы будем получать величины напряжений, соответствующие моментам времени t_n^* , что эквивалентно искажению формы преобразуемого сигнала. Это искажение можно трактовать как паразитную модуляцию фазы входного сигнала некоторой функцией времени $\varphi(t)$, основная частота которой равна частоте входного сигнала. При этом величина периода, а следовательно, и основная частота преобразуемого сигнала не изменяется.

Очевидно, что максимальному фазовому искажению подвергается точка a (рис. 10), для которой

$$\Delta \varphi = \Delta \varphi_{max} = \frac{\Delta t}{T} = f(t_n^* - t_n), \quad (9)$$

где $\Delta \varphi_{max}$ - максимальное искажение фазы, выраженное в долях периода входного сигнала (при условии максимальной амплитуды U_{bx}).

Все остальные точки будут сдвинуты на угол $\Delta \varphi = 0 \div \Delta \varphi_{max}$,

$$\Delta \varphi_{cp} = \frac{\Delta \varphi_{max}}{2} \quad (10)$$

$\Delta \varphi_{cp}$ всегда можно учесть в процессе анализа, поэтому будем считать, что $\Delta \varphi^* = \left(-\frac{\Delta \varphi_{max}}{2}\right) \div \left(+\frac{\Delta \varphi_{max}}{2}\right)$.

$$\text{Тогда } |\Delta \varphi^*|_{max} = \frac{\Delta \varphi_{max}}{2}$$

Зависимость $|\Delta \varphi^*|_{max}$ от частоты при условии, что входной сигнал имеет предельный спектр, приведена на рис. 12.

Известно [7], что сдвиг фаз звуковых сигналов в диапазоне выше 500 гц не влияет на разборчивость речи. На частотах вплоть до 500 гц преобразователь вносит искажения фазы не более 0,5%, чем практически можно пренебречь. На более высоких частотах искажения фазы возрастают, но это не имеет значения для восприятия речи.

Приведенные рассуждения справедливы для всех случаев, когда отсутствует запоминание входного сигнала на время всего цикла преобразования.

В заключение следует отметить, что данный преобразователь может быть использован для ввода в ЭВМ не только звуков речи, но и других сигналов, если погрешности преобразования, найденные по описанной выше методике, не превосходят для этих сигналов допустимую величину.

Автор выражает признательность Долгиву Т.Б., Зуеву А.Ф. и Милину А.И. за консультации и помощь при разработке преобразователя.

Л и т е р а т у р а

1. Загоруйко Н.Г. Об обмене устной информацией между человеком и вычислительными системами (данный сборник).
2. Котельников В.А. О пропускной способности эфира и проволоки в электросвязи. Изд. ред. упр. связи РККА, 1933.
3. Агеев Д.В. Активная полоса частот спектра функции времени. Труды Горьковского политехнического ин-та II, №1, 1955.
4. Харкевич А.А. Спектры и анализ. ФМ, 1962.
5. Волошин Г.Я. Спектральный анализ речевых сигналов с помощью ЭВМ (данный сборник).
6. Зайцев Т.Е. Об исследовании характеристик речевого процесса с помощью динамического спектрографа. Электро-связь № 10, 1958.
7. Флетчер Х. Речь и слух в связи. Нью-Йорк, 1953.

Рис. 1

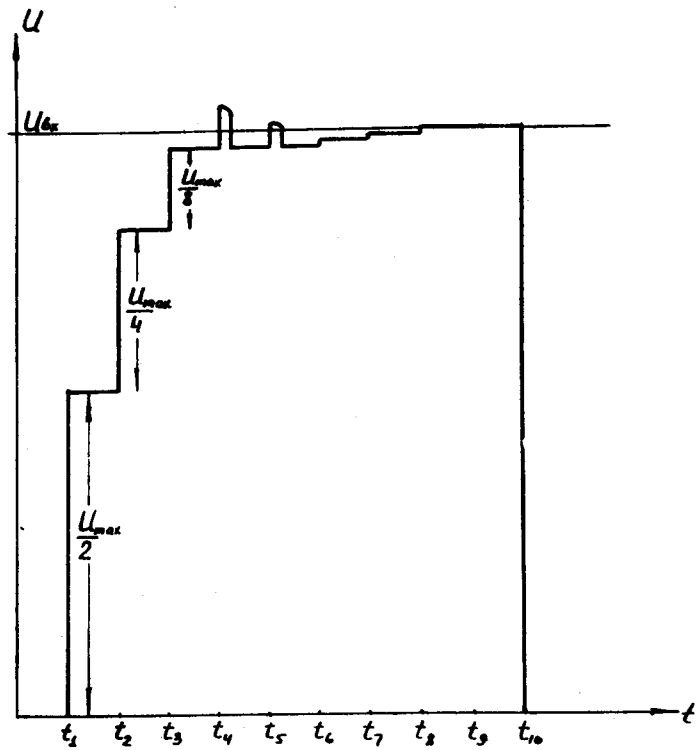
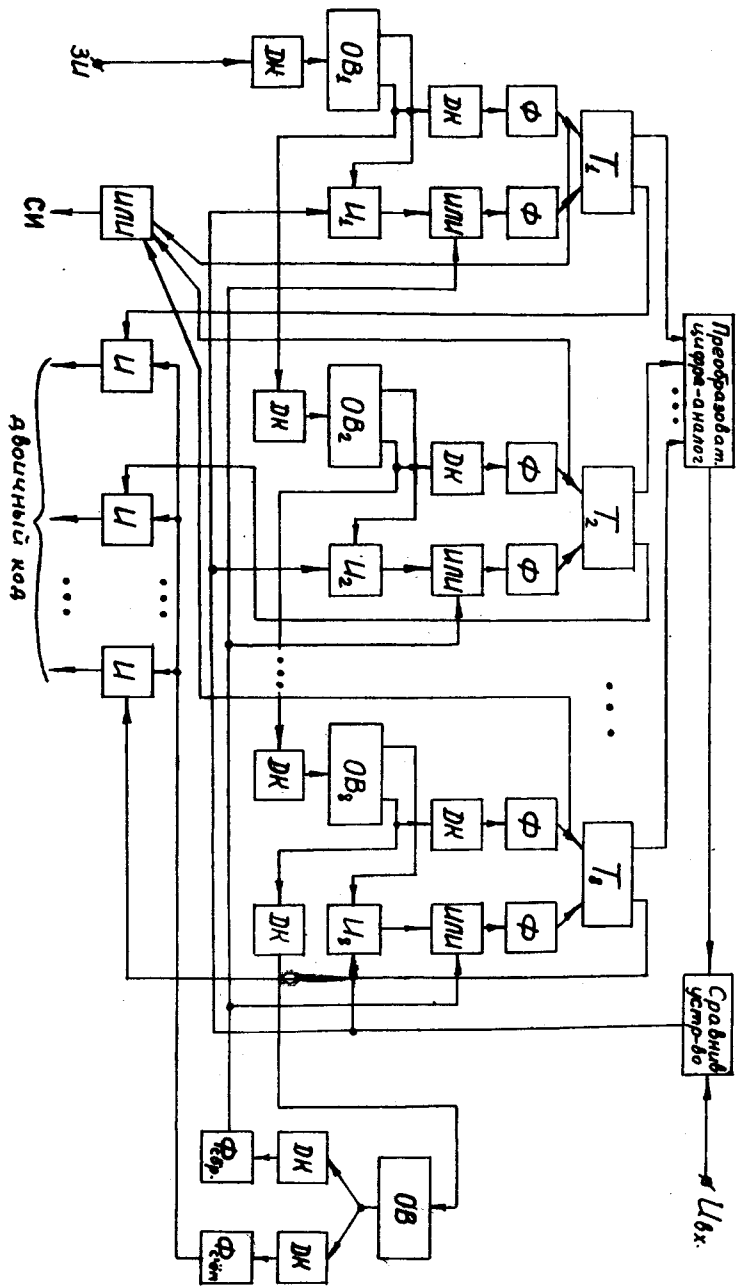


Рис. 2

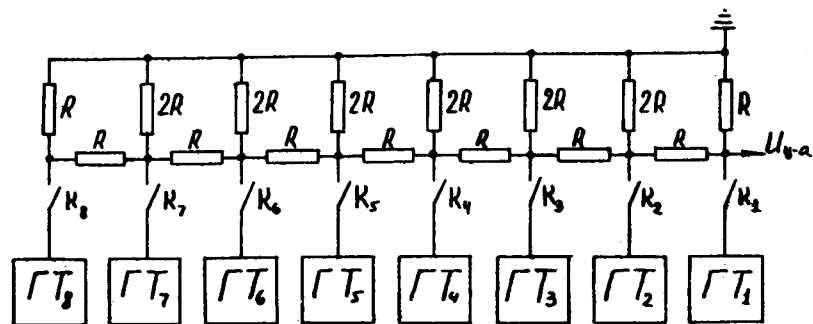


Рис. 3

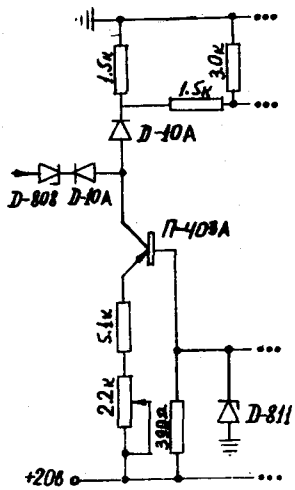


Рис. 4

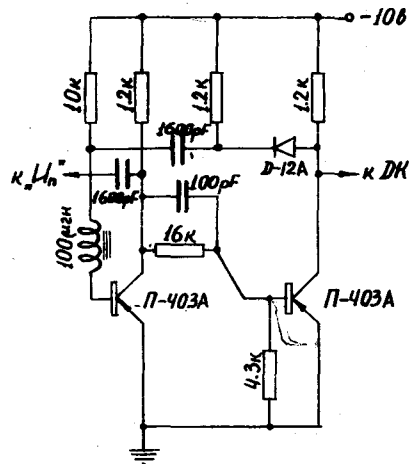


Рис. 5

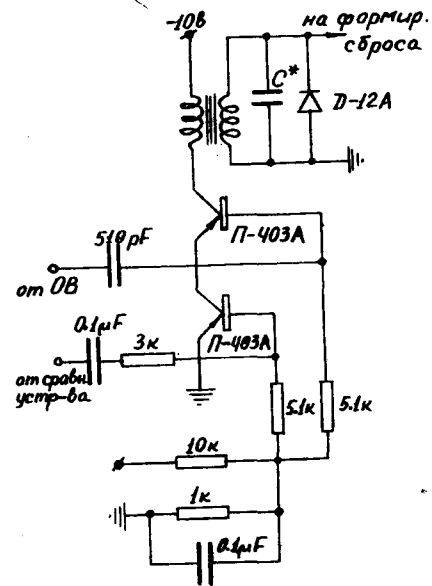


Рис. 7

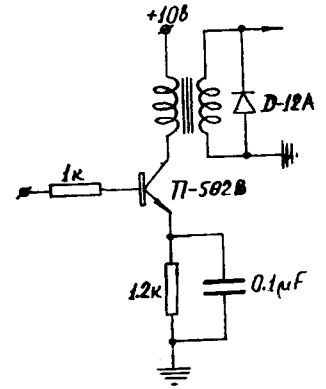


Рис. 8

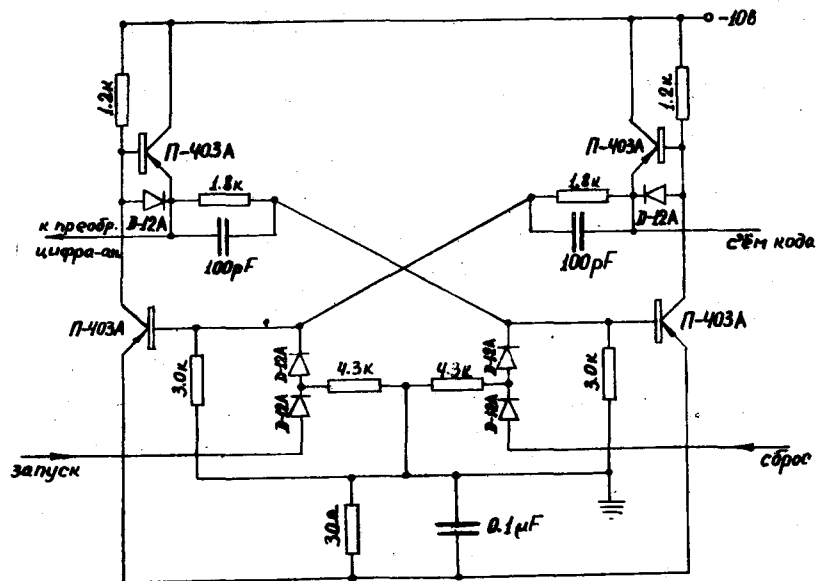


Рис. 6

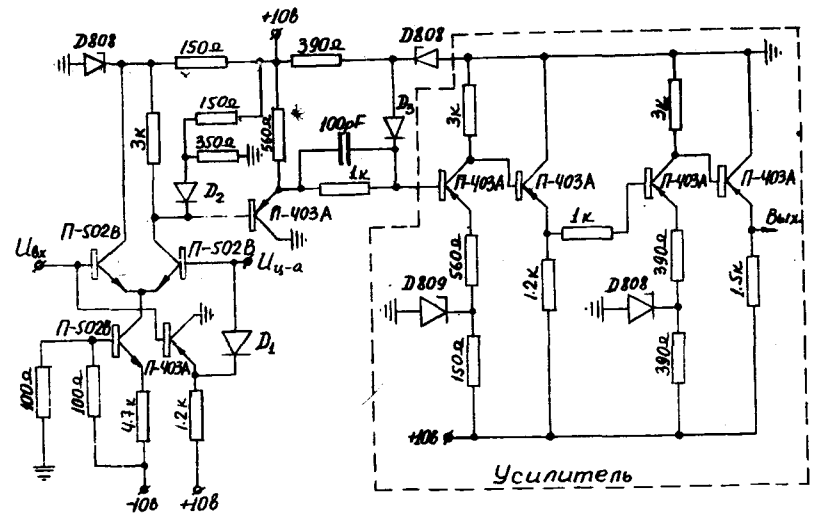


Рис. 9

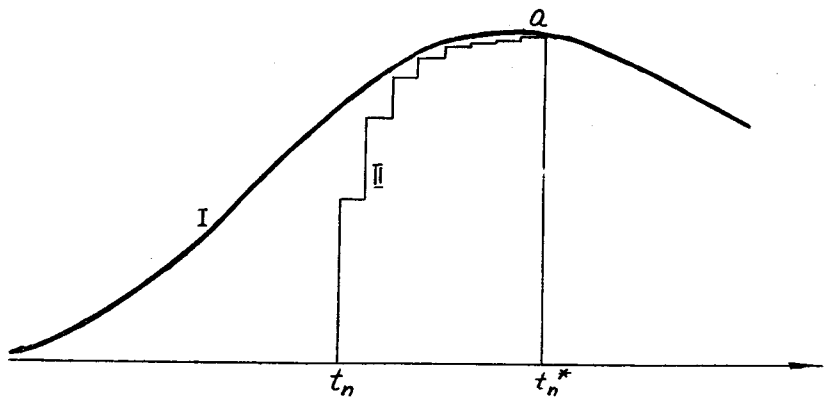


Рис. 10

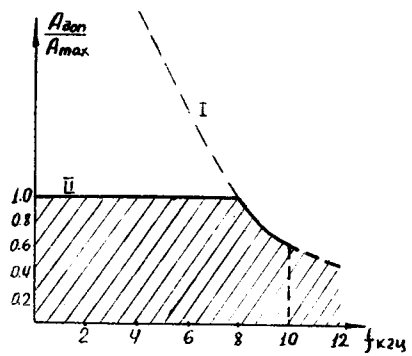


Рис. 11

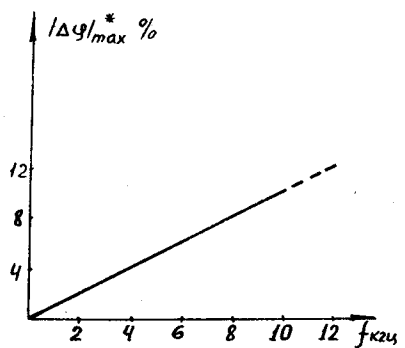


Рис. 12