

Задание для группы ПКС-21 на 24.04.2021:

Составить конспект по теме «Работа со звуком», отвечая на вопросы.

Zhernova.rita@Yandex.ru

- 1.Какая наука занимается изучением слуха и восприятием звука?
- 2.Что требуется для повышения качества воспроизведения звука?
3. Чем определяются требования к телефонам, микрофонам, громкоговорителям, к аппаратуре записи и воспроизведения звука?
4. В каких единицах измеряется высота тона?
5. Что называется *постоянной времени слуха*?
6. Единица измерения уровня громкости.
7. Что называется громкостью?
8. В каких единицах измеряется громкость?
9. Приведите примеры динамических диапазонов для некоторых видов звуковых сигналов.
- 10.Объясните различия первичных и вторичных звуковых сигналов.
- 11.Как можно охарактеризовать понятие тембра?

Работа со звуком

Основные свойства слуха

Изучением слуха и восприятия звука занимается наука под названием психоакустика. В этой науке на основе субъективных наблюдений устанавливаются закономерности и взаимосвязи между объективными характеристиками звука и ощущениями его восприятия. Эти взаимосвязи лишь иногда описываются с помощью эмпирических формул, чаще представляются в виде графиков, а иногда они носят лишь описательный характер, например, для тембра звука.

Изучение психоакустики радиоинженерами, и особенно специалистами в области звукотехники, в настоящее время необходимо потому, что результаты исследований в этой области используются в целом ряде разделов звукотехники.

В области аналоговой звукотехники результаты, полученные в психоакустике, позволяют разрабатывать регуляторы уровня и тембра, акустические головки и акустические системы, шумоподавители, эквалайзеры и усилители мощности с учетом слухового восприятия. Без знаний в области психоакустики нельзя понять, почему мы не слышим 20-процентные нелинейные искажения громкоговорителей на низких частотах и почему нас крайне раздражают нелинейные искажения менее 0,1%, возникающие в транзисторных усилителях.

В цифровой звукотехнике до сих пор существует необходимость понижения шума квантования, возникающего при аналого-цифровом преобразовании. Этот шум действует на слух значительно более раздражающе, чем шум аналоговой аудиоаппаратуры. Хотя максимальное отношение сигнал-шум для цифровой аудиоаппаратуры достигает 96-120 дБ, при воспроизведении записи реальных

музыкальных программ оно не превышает 50-70 дБ. Существенное снижение шума квантования при разработке цифровой аудиоаппаратуры также возможно только с учетом особенностей слуха человека.

В системах связи и радиовещания крайне ограничены возможности выбора свободных частотных диапазонов, поэтому остро стоит задача понижения скорости цифровых звуковых потоков без снижения субъективного качества звучания.

В области цифровой звукозаписи для повышения качества воспроизведения звука требуется повышение частоты дискретизации и числа разрядов без увеличения размеров носителя и сокращения времени записи. Для этого требуется осуществить значительное сжатие звука с уменьшением скорости цифрового потока в 4-10 раз. Поэтому в этих областях техники стали использовать субполосное кодирование, при котором звуковой звук разделяется на множество субполос шириной близкой к критической полосе слуха, а кодирование осуществляется в каждой полосе отдельно с выбором числа разрядов так, чтобы шумы квантования не ощущались на слух. Было разработано множество систем сжатия цифровых аудиоданных, основанных на различных моделях слухового восприятия, таких как MASCAM, MUSICAM, ATRAC, ASPEC и других.

В настоящее время такое кодирование звука используется в европейских цифровых системах радиовещания DAB, DRM, американской Dolby AC-3, оптических дисках системы DVD-Audio, магнитооптических минидисках, звуковых дисках системы MP-3, в сети Internet для передачи высококачественного звука.

Стереофонические звуковые системы строятся на основе знаний бинауральных особенностей слуха человека. Такая способность слуха позволяет воспринимать объемное звучание с локализацией источников звука в пространстве.

Восприятие по частоте

Представляет значительный интерес частотная избирательность слухового анализатора, так как от этого параметра зависят требования к электроакустической аппаратуре. Для количественной оценки избирательных свойств слуха удобно воспользоваться понятием высоты звука. Возможность определения высоты звука является важнейшим свойством слуховой системы. Это свойство имеет огромное значение для идентификации и классификации звуков в окружающем звуковом пространстве, эта же способность слуховой системы лежит в основе восприятия интонационного аспекта музыки, т. е. мелодии и гармонии. В соответствии с международным стандартом ANSI-994 **«Высота (Pitch) - это атрибут слухового ощущения в терминах, в которых звуки можно расположить по шкале от низких к высоким. Высота зависит главным образом от частоты звукового стимула, но она также зависит от звукового давления и от формы волны»**. Таким образом, высота - это линейная классификация звуковых сигналов, в отличие от громкости, о которой можно только сказать больше или меньше, т. е. это - относительная классификация.

Прежде всего, необходимо отметить, что слуховая система способна различать высоту звука только у периодических сигналов, поэтому определяющим параметром для различения высоты тона является частота сигнала. Если это сложный звук, то высоту слуховая система может присвоить по его основному тону, но только если он имеет периодическую структуру, т. е. его спектр состоит из гармоник (обертонов,

частоты которых находятся в целочисленных отношениях). Если это условие не выполняется, то определить высоту тона слуховая система не может. Например, звуки таких инструментов как тарелки, гонги и др. не имеют определенной высоты.

Изучение связи частоты звука и воспринимаемой высоты предпринималось еще Пифагором, а также многими известными физиками: Галилеем, Гельмгольцем, Омом и др. В настоящее время на основе тщательных экспериментов, в процессе которых слушателю предъявлялись два звука разной частоты с просьбой расположить их по высоте, установлена зависимость высоты тона от частоты сигнала, показанная на рис. 1.

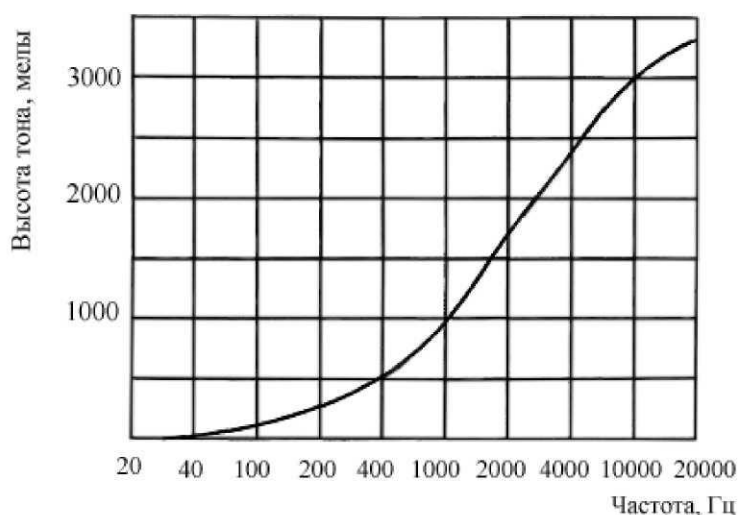


Рис. 1. Связь частоты и высоты тона

Высота тона измеряется в специальных единицах - мелах. Один мел равен ощущаемой высоте звука частотой 1000 Гц при уровне 40 дБ. Как видно из рисунка, эта связь не линейна - при увеличении частоты, например, в три раза (от 1000 до 3000 Гц), высота повышается только в два раза (от 1000 до 2000 мел). Нелинейная зависимость особенно выражена на низких и высоких частотах. В средней части диапазона частот изменение высоты тона в мелах пропорционально логарифму частоты.

Ощущение высоты чистого тона связано не только с частотой, но и с интенсивностью звука и его длительностью. Как показали различные исследования, при повышении интенсивности звука громкие низкие звуки кажутся еще ниже, а высокие звуки с повышением громкости кажутся слегка выше (только для средних частот 1-2 кГц влияние интенсивности незаметно). Следует отметить, что эта зависимость весьма незначительна, а для сложных музыкальных звуков почти незаметна.

Большая точность слуха по частоте предъявляет довольно жесткие требования к точности хода лентопротяжных механизмов магнитофонов и к точности скорости и стабильности вращения дисков с записью звука.

Уровень громкости и громкость

Между тем человек может довольно точно установить равенство громкостей двух звуков любого частотного состава. Это свойство слуха было использовано для введения понятия уровня громкости, позволяющего определить громкость одного звука относительно другого, принятого в качестве опорного.

За единицу уровня громкости приняли величину уровня интенсивности чистого тона с частотой 1000 Гц. Единица уровня громкости называется *фон*. Численно уровень громкости в фонах равен уровню интенсивности звука в децибелах на частоте 1000 Гц.

Громкостью называется субъективное ощущение, позволяющее слуховой системе располагать звуки по шкале от тихих до громких звуков. Громкость звука связана, прежде всего, со звуковым давлением. Обычно, чем больше звуковое давление, тем громче звучит акустическая система. Однако это не всегда так. Громкость также зависит от частоты, спектрального состава, длительности звука и его локализации в пространстве. Термин громкость иногда заменяют абсолютной громкостью, относительной или субъективной громкостью – это все одно и то же.

Для оценки громкости принята специальная единица – *сон*. Громкость 1 сон – это громкость тонального звука с частотой 1 кГц с уровнем 40 дБ.

Зависимость громкости от уровня звукового давления является сугубо нелинейной, у нее логарифмический характер. При увеличении уровня давления звука на 10 дБ его громкость возрастает в 2 раза. Это значит, что уровням громкости 40, 50 и 60 фон соответствуют громкости 1, 2 и 4 сона.

Чтобы представить громкость звуков, с которыми мы постоянно сталкиваемся в жизни, в таблице 2.1 приведены уровни громкости и громкость для наиболее типовых звучаний.

Таблица 1

Средний уровень громкости и громкость наиболее часто встречающихся звуков и шумов

Источник шума и место его измерения	Уровень громкости, фон	Громкость, сон
Шумы вне помещений		
Авиационный мотор на расстоянии 5 м	116-120	346-556
Громкий автомобильный сигнал на расстоянии 8 м	95- 100	57-88
Электropоезд на эстакаде на расстоянии 6 м	90	38
Шум в поезде метро во время движения	85-90	25-38
Автобус (полный ход) на расстоянии 5 м	85-88	25-32,2
Трамвай на расстоянии 10-20 м	80-85	17,1-25

Источник шума и место его измерения	Уровень громкости, фон	Громкость, сон
Грузовой автомобиль в городе на расстоянии 5-20 м	60-75	4,35-11,4
Легковой автомобиль в городе на расстоянии 5-20 м	50-65	2,2-5,87
Свисток милиционера на расстоянии 20 м	70	7,95
Шумная улица без трамвайного движения	60-75	4,35-11,4
Обычный средний шум на улице	55-60	3,08-4,35
То же, в момент затишья днем	40	1,0
Тихая улица (без движения транспорта)	30-35	0,36-0,62
Тихий сад	20	0,097
Производственные шумы		
Удары молота по стали, клепальная машина на расстоянии 2-4 м	110-113	215-288
Котельные цехи	100-103	88-116
Общий шум в ткацком цехе	96-100	62-88
Деревообрабатывающая фабрика	96-98	62-74
Театры, школы, больницы		
Симфонический оркестр	80-100	17,1-88
Аплодисменты	60-75	4,4-11,4
Громкая музыка по радио	80	17,1
Радиоцентр (студия во время исполнения)	40-50	1-2,2
Актзовые залы в школах во время перерыва	55-62	3,1-4,7
Библиотеки	25-30	0,2-0,36
Больницы	20-30	0,1-0,36
Шумное собрание	65-70	5,87-7,95
Громкий разговор на расстоянии 1 м	65-70	5,87-7,95
Обычный на расстоянии 1 м	55-60	3,08-4,35
Жилые помещения		
Разговор трех человек в обычной комнате	45-50	1,5-2,2
Шепот средней громкости на расстоянии 1 м	20	ОД
Степени музыкальной громкости		
Форте-фортиссимо	100	88
Фортиссимо	90	38
Форте	80	17,1
Меццо-форте	70	7,95
Меццо-пиано	60	4,35
Пиано	50	2,2
Пианиссимо	40	1,0
Пиано-пианиссимо	30	0,36

Слуховой аппарат инерционен: при исчезновении звука слуховое ощущение исчезает не сразу, а постепенно, уменьшаясь до нуля.

Время, в течение которого ощущение по уровню громкости уменьшается на 9-10 фон, называется *постоянной времени слуха*. В среднем она равна 30-50 мс.

ЗВУКОВЫЕ СИГНАЛЫ

Определения

Различают **первичные** и **вторичные** звуковые сигналы. К первичным относятся: сигналы, создаваемые музыкальными инструментами, пением, речью; шумовые сигналы, создаваемые для сопровождения различных музыкальных и речевых художественных передач (шум поезда, рокот моря, свист ветра и т. п.).

При оценке трактов вещания и связи полагают, что каждый акустический сигнал почти всегда является случайным в вероятностном смысле и несет в себе информацию, соответствующую его объему. Строго говоря, звуковые сигналы художественных программ, не могут рассматриваться как случайные в вероятностном смысле, так как они могут быть хорошо известны слушателям, и поэтому не нести в себе информацию. Для слушателя эти сигналы служат средством удовлетворения эстетических потребностей, а не получения информации. Однако при расчете каналов звукопередачи эти сигналы считают несущими информацию, соответствующую их объему. Хотя в музыкальных сигналах очень большие участки могут иметь периодический характер, в среднем, для больших интервалов времени их также можно рассматривать как случайные. Поэтому параметры звуковых сигналов определяются их распределениями по уровню, по частотному диапазону и во времени¹.

Ко вторичным звуковым сигналам относятся сигналы, воспроизводимые электроакустическими устройствами, то есть первичные сигналы, прошедшие по электроакустическим трактам связи.

Динамический диапазон

В процессе любой передачи уровень акустического сигнала непрерывно изменяется, причем диапазон его изменения может быть довольно широким.

Динамический диапазон для некоторых видов звуковых сигналов приведен в таблице 1.

Таблица 1 Динамический диапазон для некоторых видов звуковых сигналов

Вид сигнала	Динамический диапазон, дБ
Речь диктора	25-35
Художественное чтение	35-45
Телефонные разговоры	35-45
Небольшие ансамбли	45-55
Симфонический оркестр	65-75
Рок-музыка	до 118

Как видно из таблицы, для передачи натурального динамического диапазона требуется высококачественная аппаратура (с большим запасом линейной

части амплитудной характеристики). В большинстве случаев динамический диапазон первичных звуковых сигналов превышает возможности аналоговых средств связи и вещания. Поэтому при их использовании приходится предварительно сжимать динамический диапазон или же мириться с появлением значительных искажений в тракте передачи.

Средний уровень

Средний уровень интенсивности акустического сигнала можно определять или по слуховому ощущению (субъективное среднее), или как средний статистический по интенсивности для длительных интервалов времени (среднее длительное), или как средний, измеряемый прибором, имеющим небольшую постоянную времени (объективное среднее). Для вторичных сигналов достаточно определять только средний уровень по ощущению, для первичных - необходимо знать все средние уровни.

Частотный диапазон и спектры

Звуковой сигнал от каждого из первичных источников звука, как правило, имеет непрерывно изменяющуюся форму и состав спектра. Спектры могут быть высоко- и низкочастотными, дискретными и сплошными. У каждого источника звука, даже того же самого типа (например, скрипка в оркестре), спектры имеют индивидуальные особенности, что придает звучанию характерную окраску. Эту окраску называют **тембром**. Существуют понятия тембра скрипки, тромбона, органа и т. п., а также тембра голоса: звонкий, когда подчеркнуты высокочастотные составляющие; глухой, когда они подавлены. В первую очередь, представляют интерес: средний спектр для источников звука каждого типа, а для оценки искажений сигнала - спектр, усредненный за длительный интервал времени (15 с для информационных сигналов и 1 мин - для художественных). Усредненный спектр является, как правило, сплошным и достаточно сглаженным по форме.

Сплошные спектры характеризуются зависимостью **спектральной плотности** от частоты (эту зависимость называют энергетическим спектром). Спектральной плотностью называется интенсивность звука в полосе частот шириной, равной единице частоты. Для акустики эту полосу берут равной 1 Гц.

Временные характеристики акустического сигнала

К временным характеристикам сигнала относятся уровнеграмма и время корреляции. Уровнеграмма сигнала дает возможность определить резкие переходы интенсивности и, следовательно, с ее помощью можно предъявить требование к постоянным времени трактов передачи сигнала. Такие временные характеристики сигнала, как время корреляции, используют редко, хотя опыты показывают, что этот параметр играет значительную роль при определении качества звучания.

Первичный речевой сигнал

Каждому человеку свойственна своя манера произносить звуки речи (своего рода устный почерк). Произношение звуков речи зависит, например, от удара, соседних звуков и других факторов. Физические реализации ограниченного числа обобщенных звуков речи называются **фонемами**. Фонема - это то, что человек хочет

произнести, а звук речи - это то, что человек фактически произносит. Фонема по отношению к звуку речи играет ту же роль, что и образцовая буква, называемая **графемой** (например, печатная курсивом) по отношению к ее рукописной форме в конкретном написании.

В русском языке насчитываются 41 основная и 3 неясно звучащих фонемы: 6 гласных (а, о, у, э, и, ы), 1 полугласная (й) и 34 согласных. Гласные буквы я, ю, е, е соответствуют или составным фонемам: йа, йу, йо, йэ, или служат для смягчения предыдущей согласной. Согласных фонем больше, чем согласных букв, так как ряд согласных букв соответствует двум фонемам: мягкой и твердой. Только твердых фонем 3 (ш, ж, ц), только мягких - 1 (ч). Остальные 15 существуют в обоих видах: твердом и мягком.

Импульсы потока воздуха, создаваемые голосовыми связками с достаточной точностью, могут считаться периодическими. Соответствующий период повторения импульсов называют периодом основного тона голоса T_0 . Обратную величину $F_0 = 1/T$ называют частотой основного тона. Если связки тонкие и сильно напряжены, то период получается коротким и частота основного тона - высокой; для толстых, слабонапряженных связок частота основного тона низкая. Эта частота для всех голосов лежит в пределах от 70 до 450 Гц. При произнесении речи она непрерывно изменяется в соответствии с ударением и подчеркиванием звуков и слов, а также для проявления эмоций (вопрос, восклицание, удивление и т. д.). Изменение частоты основного тона называют **интонацией**. У каждого человека свой диапазон изменения частоты основного тона (обычно он бывает немногим более октавы) и своя интонация. Последняя имеет большое значение для узнаваемости говорящего. Основной тон, интонация, устный «почерк» и тембр (окраска) голоса могут служить для опознавания человека. При этом степень достоверности опознавания выше, чем по отпечаткам пальцев. Это свойство используют в разработанной в последнее время аппаратуре, реагирующей только на определенные голоса.

Известно, что для передачи одного и того же сообщения по телеграфу и по речевому тракту требуется различная пропускная способность тракта: для телеграфного сообщения не более 100 бит/с, а для речевого - около 100000 бит/с (полоса равна 7000 Гц, динамический диапазон 42 дБ, то есть требуется семизначный код, откуда имеем: $2 \cdot 7000 \cdot 7 = 98000$ бит/с), то есть в 1000 раз большая.

Может показаться, что речевой сигнал имеет огромную избыточность. Это неверно и вот почему. Образование звуков речи происходит путем подачи команд к мускулам артикуляционных органов речи от речевого центра мозга. Этих сигналов не более десяти, при этом они изменяются медленно (в темпе смены звуков речи, то есть не более десяти раз в секунду). Поэтому общий поток сообщений для них составляет около 100 бит/с.

Вторичный сигнал

В идеальном случае вторичный сигнал должен точно воспроизводить первичный, но это не всегда требуется, так как слух человека может и не заметить их несоответствие. К тому же на практике их точное соответствие часто невозможно или очень трудно осуществить. При художественном вещании, телевидении и звукозаписи надо стремиться к этому соответствию в пределах, при которых слуховое ощущение, создающееся у слушателя, было бы близко к тому ощущению, которое он получает, находясь в месте исполнения данной программы при создании достаточно

хороших акустических условий. Для информационных программ вещания и телефонной связи этого соответствия добиваются в первую очередь для получения полной понятности речи, а затем для достаточно высокого качества звучания. Только в этом случае необходимо стремиться к более точному соответствию вторичного сигнала первичному. В обоих случаях существенную роль играют экономические соображения.

Все несоответствия первичного и вторичного сигналов являются искажениями в широком смысле этого понятия. Но обычно под этим термином понимают более узкий тип искажений. К ним относятся линейные, нелинейные, параметрические и переходные (временные) искажения.

Цифровое представление звуковых сигналов

Аналого-цифровое преобразование

Преобразование аналогового звукового сигнала в цифровой включает в себя несколько этапов. Сначала аналоговый звуковой сигнал подается на аналоговый фильтр, который ограничивает полосу частот сигнала и устраняет помехи и шумы. Затем из аналогового сигнала с помощью схемы выборки/хранения выделяются отсчеты: с определенной периодичностью осуществляется запоминание мгновенного уровня аналогового сигнала. Далее отсчеты поступают в аналого-цифровой преобразователь (АЦП), который преобразует мгновенное значение каждого отсчета в цифровой код или числа. Полученная последовательность бит цифрового кода, собственно, и является звуковым сигналом в цифровой форме. В результате преобразования непрерывный аналоговый звуковой сигнал превращается в цифровой – дискретный как по времени, так и по величине. Для примера на рис. 1 показана структурная схема канала цифровой записи звука.

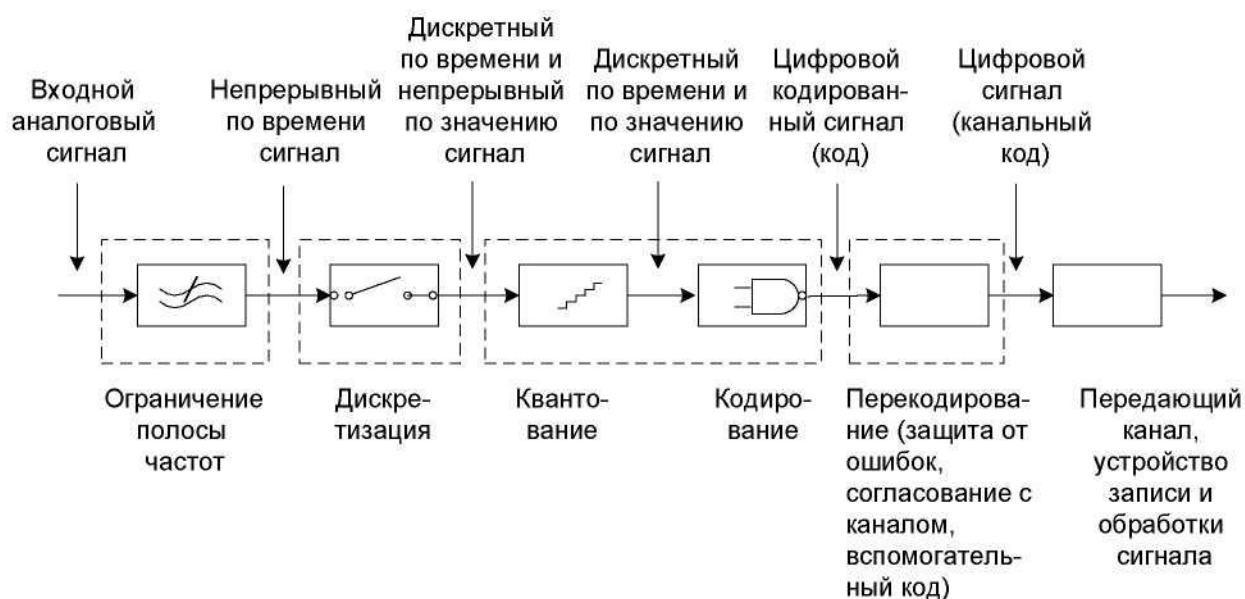


Рис. 1. Обобщенная структурная схема канала цифровой записи звука

Главную роль в процессе преобразования сигнала из аналоговой формы в цифровую играет АЦП (Analog/Digital Converter – ADC). Обратный процесс преобразование цифрового звукового сигнала в аналоговый реализуется с помощью цифро-аналогового преобразователя – ЦАП (Digital/ Analog Converter – DAC).

Затраты памяти на запись звука в цифровой форме. Зная разрядность АЦП, а точнее, количество разрядов, используемых для записи звукового сигнала в цифровой форме, можно получить некоторые интересные цифры.

Например, если умножить число разрядов кодового слова на частоту дискретизации сигнала, выраженную в герцах, то получим скорость передачи данных, которую должен обеспечивать цифровой канал записи/воспроизведения звука. Если полученную скорость передачи данных умножить на общую длительность звукового сигнала в секундах, получим объем памяти на магнитном носителе, например, на жестком диске, который потребуется для хранения звуковых данных. В случае записи стереосигнала, когда запись идет по двум (левому и правому) стереоканалам, скорость передачи данных и необходимый объем памяти удваиваются.

Цифро-аналоговое преобразование

Для воспроизведения звукового сигнала, записанного в цифровой форме, необходимо преобразовать его в аналоговую форму, то есть осуществить цифро-аналоговое преобразование сигнала (рис. 2).

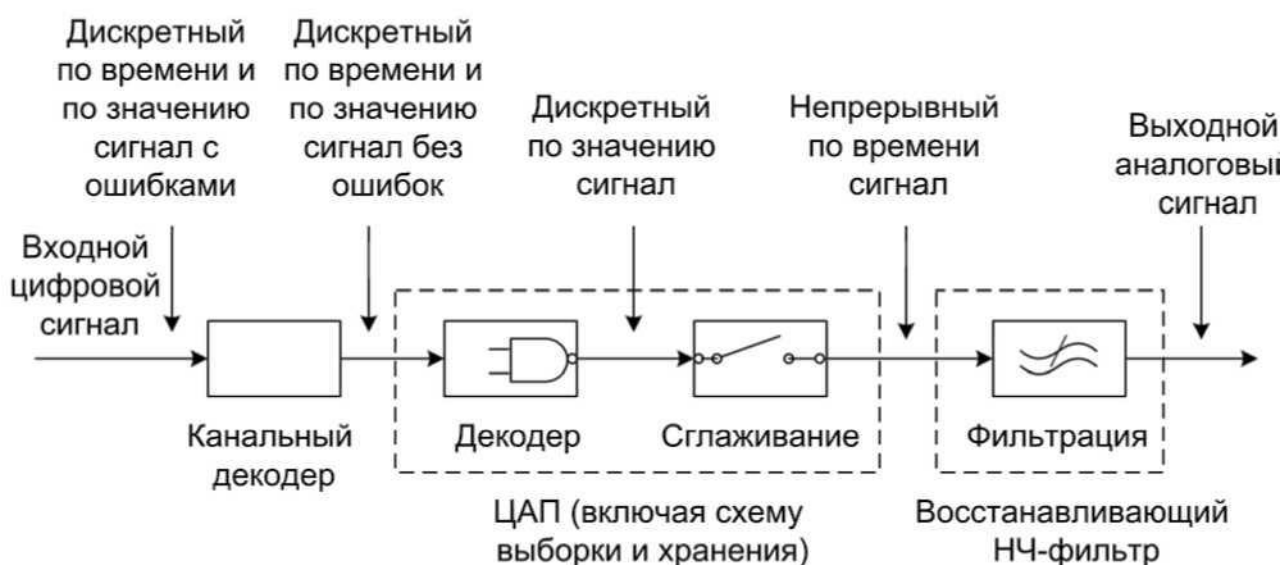


Рис. 2. Обобщенная схема преобразования цифрового сигнала в аналоговый

Цифро-аналоговое преобразование производится в два этапа.

На первом этапе из потока цифровых данных с помощью цифро-аналогового преобразователя выделяют отсчеты сигнала, следующие с частотой дискретизации. На втором этапе из дискретных отсчетов формируется путем сглаживания (интерполяции) непрерывный аналоговый сигнал. Эта операция равносильна фильтрации сигнала идеальным фильтром низкой частоты, который подавляет периодические составляющие спектра дискретизированного сигнала.

Как и в АЦП, в ЦАП широко применяется оверсэмплинг, поскольку существует проблема создания восстанавливающих (интерполирующих) аналоговых фильтров.

Полученный в результате цифро-аналогового преобразования звуковой сигнал, как правило, попадает в микшер и через линейный выход направляется в акустическую систему, в которой колебания напряжения электрического сигнала преобразуются в колебания звукового давления.

