

Министерство образования Республики Беларусь
Учреждение образования
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ

Кафедра вычислительных методов и программирования

А.И. Волковец

**Создание и обработка звука при разработке интерактивных
приложений**

Конспект лекций

для студентов специальности «Информационные системы и технологии (в
игровой индустрии)»

Минск 2018

Оглавление

Лекция 1. Основы акустики	4
Физические основы звуковых колебаний.	4
Звуковой диапазон частот	6
Динамический диапазон слуха	7
Слуховой аппарат	11
Пространственное восприятие звуков	12
Лекция 2. Преобразователи звука	15
Преобразование акустических колебаний в электрический сигнал	15
Преобразование электрического сигнала в акустические колебания	23
Акустические системы	27
Головные телефоны.....	29
Лекция 3. Аналого-цифровое преобразование.	32
Дискретизация по времени	36
Квантование по уровню	40
АЦП для цифровой обработки звука	42
Лекция 4. Цифро-аналоговое преобразование	46
Передискретизация	47
ЦАП для цифровой обработки звука	50
Лекция 5. Спектр сигнала.	59
Дискретное преобразование Фурье.	59
Обратное дискретное преобразование Фурье	68
Быстрое преобразование Фурье.....	69
Лекция 6. Фильтрация звука	70
Типы фильтров.....	71
Цифровые фильтры.....	72
Цифровая фильтрация при помощи ДПФ и ОДПФ	73
Эквалайзеры.....	76
Лекция 7. Звуковая система ПК	80
Аппаратная часть. Кодеки.....	81
Система объёмного звука Dolby Digital	86
Параметры оценки качества звукового тракта	88
Лекция 8. Кодирование звуковой информации	91
Прямой, обратный, дополнительный коды	91
Импульсно-кодовая модуляция	92
Сжатие данных.....	95
Лекция 9. Синтез музыки	100
Интерфейс MIDI.....	100
Стандарты MIDI	120

<i>Инструменты MIDI</i>	<i>124</i>
<i>MIDI-файлы</i>	<i>126</i>
<i>Лекция 10. Синтез речи</i>	<i>131</i>
<i>Параметрический синтез.</i>	<i>131</i>
<i>Компиляционный синтез.....</i>	<i>137</i>
<i>Полный синтез речи по правилам</i>	<i>137</i>
<i>Лекция 11. Обработка цифрового звука.....</i>	<i>139</i>
<i>Микширование.....</i>	<i>139</i>
<i>Мастеринг</i>	<i>141</i>
<i>Динамическая обработка.</i>	<i>143</i>
<i>Аудиоредакторы</i>	<i>150</i>

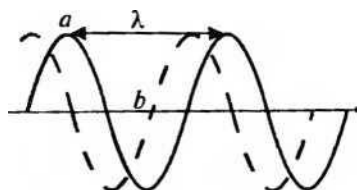
Лекция 1. Основы акустики

Звук, как явление природы, начали изучать очень давно. Так, во времена Древней Греции в IV в. до н. э., когда теория музыки только зарождалась, Архит Тарентский выдвинул революционную идею о звуке как колебании воздуха. К сожалению, он полагал, что высота звука зависит от скорости распространения волны, а не от его частоты. В буквальном переводе с греческого *akustikos* — означает «слуховой». Значение слова **акустика** — учение о звуке как упругих колебаниях, слышимых человеческим ухом.

Физические основы звуковых колебаний.

Звук в широком смысле — это волны, возникающие в воздухе (или другой упругой среде) под действием каких-либо колеблющихся предметов и создающие в ней механические колебания. *Звук в узком смысле* — субъективное восприятие колебаний в упругой среде специальным органом чувств человека и животных. Под *звуковым полем* понимают пространство упругой среды, обычно воздушной, в которой распространяются звуковые волны (колебания). Природа звуковых волн такова, что при деформации среды в каком-либо элементарном участке в соседних с ним участках возникают последовательные во времени сжатия и разряжения среды. Источниками звука могут быть, например, голосовые связки человека, струны музыкальных инструментов или любой другой вибрирующий предмет, заставляющий с определенной скоростью колебаться окружающие его частицы воздуха.

Скорость распространения звуковых волн в атмосфере при нормальных температуре и давлении близка к значению $c_{36}=340$ м/с. В средах с большой плотностью (жидких, твердых) скорость распространения соответственно повышается. В неограниченном пространстве звук распространяется в виде бегущей волны. В акустике существует понятие длина волны звукового колебания. Эта величина определяется отрезком на предполагаемой оси, расположенной в направлении распространения звука, на котором уместается полный цикл изменений звукового давления. Иначе говоря, длина волны — это наименьшее расстояние между точками звукового поля с одинаковыми фазами колебания



Длина звуковой волны λ связана с частотой колебаний F соотношением

$$\lambda = \frac{c_{зв}}{F} \quad (1.1)$$

При распространении звуковой волны в каждом из участков звукового поля наблюдаются периодические колебания частиц среды: сжатия и разрежения. Такие локальные колебания характеризуются знакопеременным смещением частиц, отклонением их от первоначального статического положения. Скорость колебательного движения частиц среды при распространении в ней звуковой волны называется **колебательной скоростью** (v). Не следует путать это понятие со скоростью распространения $c_{зв}$ звуковой волны. Колебательная скорость возрастает при повышении частоты и амплитуды акустического сигнала, однако она практически на несколько порядков меньше скорости распространения звуковой волны.

Поочередное сжатие и разрежение в данном участке поля вызывает соответственно изменение давления относительно его статического значения. В данном случае в качестве последнего выступает атмосферное звуковое давление. Разность между мгновенными значениями давления при сжатиях и разрежениях среды в данной точке относительно статического атмосферного давления называется **звуковым давлением** (p). Таким образом, звуковое давление считают положительным в фазе сжатия и отрицательным в фазе разрежения. Звуковое давление является скалярной знакопеременной величиной, измеряется в паскалях ($1 \text{ Па} = 1 \text{ Н/м}^2$). Звуковое давление гораздо меньше атмосферного (статического). Диапазон изменения $p : 10^{-5} \dots 20 \text{ Па}$.

Другим важным параметром звукового поля является **интенсивность** (или сила) **звука** (I). Под этим параметром понимают поток звуковой энергии, переносимый средой в направлении распространения через единицу поверхности, ему перпендикулярной, за единицу времени. Единица измерения силы звука — Вт/м^2 . Она связана со звуковым давлением квадратичной зависимостью. Это значит, что изменение силы звука пропорционально соответствующему изменению величины звукового давления, возведенному в квадрат (I пропорционально p^2). Объясняется это тем, что увеличение звукового давления обязательно сопровождается пропорциональным увеличением колебательной скорости (v) частиц воздуха. А сила звука зависит одновременно от этих обеих величин, т.е.

$$I = pv = k \cdot p^2. \quad (1.2)$$

Следовательно, прирост звукового давления в 2 раза влечет за собой увеличение силы звука в 4 раза и т.д.

Звуковой диапазон частот

Область акустических колебаний, способных создавать ощущение звука при воздействии на орган слуха, ограничена по частоте. Для большинства людей 18—25 лет, обладающих нормальным слухом, полоса частот колебаний, воспринимаемых в виде звука, находится в пределах от 16 ... 20 Гц и до 18 ... 20 000 Гц. Эту полосу частот принято называть звуковым диапазоном, а частоты, лежащие в его пределах, — **звуковыми частотами**.

С нижней границей звукового диапазона граничит диапазон инфразвуковых частот, воздействие которых на человека считают вредным, так как они могут вызывать неприятные ощущения с серьезными последствиями. В природе инфразвуковые колебания могут возникать при волнениях в море, колебаниях земной среды и пр.

Выше звукового диапазона располагается диапазон ультразвуковых механических колебаний. Ультразвук широко используется в радиоэлектронике для создания гидролокаторов (сонаров), эхолотов, устройств, служащих для обработки радиотехнических сигналов, например, фильтров, линий задержки, преобразователей формы сигналов (в миниатюрном исполнении с использованием принципа поверхностных акустических волн — ПАВ), а так же в медицине и промышленности. Частота звуковых колебаний определяет высоту (тон) звука: самые медленные колебания воспринимаются как низкие, басовые ноты; самые быстрые — как высокие звуки, напоминающие, например, комариный писк. Не все люди и не всегда хорошо слышат все частоты звукового диапазона. С возрастом верхняя граница слышимых частот значительно понижается.

В музыкальной акустике принято делить частотный диапазон на октавы и доли октавы. Понятие октава соответствует изменению частоты F в два раза; весь диапазон звуковых частот охватывается 10 октавами. Музыкальная шкала октавы подразделяется на 12 полутонов, что соответствует приращению частоты в ΔF раз:

$$\Delta F = \sqrt[12]{2} = 1,054$$

или тонам звуков двух смежных клавиш рояля.

Динамический диапазон слуха

Человеческий слух улавливает не все звуки. Человек начинает слышать при силе звука, превышающей или равной некоторой величине, называемой **порогом слышимости** (или слуховым порогом). Более слабые звуки слухового ощущения не вызывают. С увеличением силы звука достигается нормальная слышимость, а затем при еще больших амплитудах звуковых колебаний к воспринимаемому звуку добавляется осязаемое ощущение давления, а дальнейший рост силы звука вызывает болезненное раздражение органа слуха. Так называемый **болевой порог** ограничивает область слышимости при высоких уровнях интенсивности. Чувствительность человеческого уха зависит от частоты приходящего сигнала, поэтому уровень порога слышимости для разных частот различный. На рис. 1.1 приведены значения силы звука и звукового давления, при которых звуковые сигналы с различными частотами становятся едва слышимыми. Человеческое ухо наиболее чувствительно к частотам от 1000 до 5000 Гц.

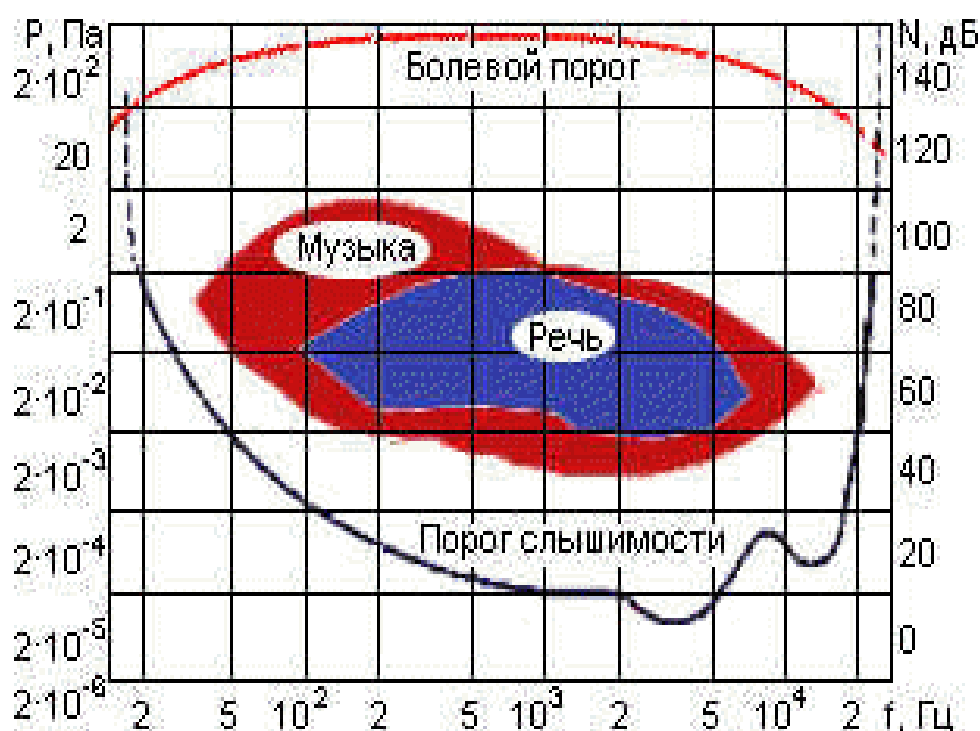


Рис. 1.1 Области слышимости

Согласно психофизическому закону Вебера—Фехнера, слух одинаково оценивает равные **относительные** изменения силы звука. Другими словами, изменение громкости кажется человеку одинаковым, если сила звука изменилась в одно и то же число раз (или на один и тот же процент относительно своей первоначальной

величины), при этом восприятие не зависит от абсолютного уровня силы звука. Так, двукратный рост уровня тихого и громкого звука воспринимаются одинаково, хотя абсолютные приращения | звукового давления существенно различны. Это объясняется логарифмическим законом восприятия. Наши ощущения изменений громкости пропорциональны не изменениям силы звука, а логарифму этих величин:

$$L = C \cdot \log_{10} \frac{I_2}{I_1} \quad (1.3)$$

где L — кажущееся изменение громкости;

I_1, I_2 — сила звука соответственно до и после его изменения;

C — коэффициент пропорциональности.

Поэтому увеличение или уменьшение силы звука принято измерять в специальных логарифмических единицах. Причем каждое десятикратное изменение силы звука (звуковой энергии) оценивается единицей, называемой *Бел* (Б). На практике используется в основном производная от Бела единица измерения, равная десятой части Бела, т.е. децибел (дБ). Изменение уровня силы звука, выраженное в дБ, равно численному значению десятичного логарифма отношения сравниваемых уровней, умноженному на 10, т.е.

$$N = 10 \log_{10} \frac{I_2}{I_1} = 10 \log_{10} \frac{k \cdot p_2^2}{k \cdot p_1^2} = 20 \log_{10} \frac{p_2}{p_1} \quad (1.4)$$

Пользуясь децибелами для оценки различия в уровнях, мы можем теперь указать меру некоторых из упомянутых ранее величин

Таблица 1

Наименование измеряемой величины	Отношение по силе звука	Отношение по звуковым давлением	Отношение в децибелах
Минимальный перепад уровней звука, улавливаемый слухом	1,25 : 1	1,12 : 1	1
Динамический диапазон слуха	10^{12}	10^8	120
Натуральный динамический диапазон симфонического оркестра	10^8	10^4	80

Теперь о том, что подразумевается под понятием *громкость*. **Громкостью** называют субъективное ощущение звука, возникающее у слушателя под воздействием звуковых

колебаний. Громкость не может быть определена только величиной силы звука, так как она зависит и от частотного состава звукового сигнала, от условий его восприятия и длительности воздействия. В акустике для количественной оценки громкости используют метод субъективного сравнения измеряемого звука с эталонным, в качестве которого применяется синусоидальный тон частоты 1000 Гц. В процессе сравнения уровень эталонного тона изменяют до тех пор, пока эталонный и измеряемый звуки станут казаться равногромкими.

Итак, за **уровень громкости** принимают величину, численно равную уровню эталонного тона частоты 1000 Гц, равногромкого с данным звуком. Уровень громкости выражается в фонах. Единицы измерения уровня громкости (**фоны**) численно совпадают с уровнем звука, выраженным в децибелах, на частоте эталонного тона 1000 Гц.

Можно нарисовать семейство кривых для разных уровней звука таким образом, чтобы на точки каждой из этих кривых на различных частотах соответствовали бы одному и тому же уровню громкости (рис. 1.2).

По этим кривым, называемым **кривыми равной громкости**, или кривыми Флетчера—Менсона, можно по уровню звука на любой частоте определить соответствующий этому звуку уровень громкости.

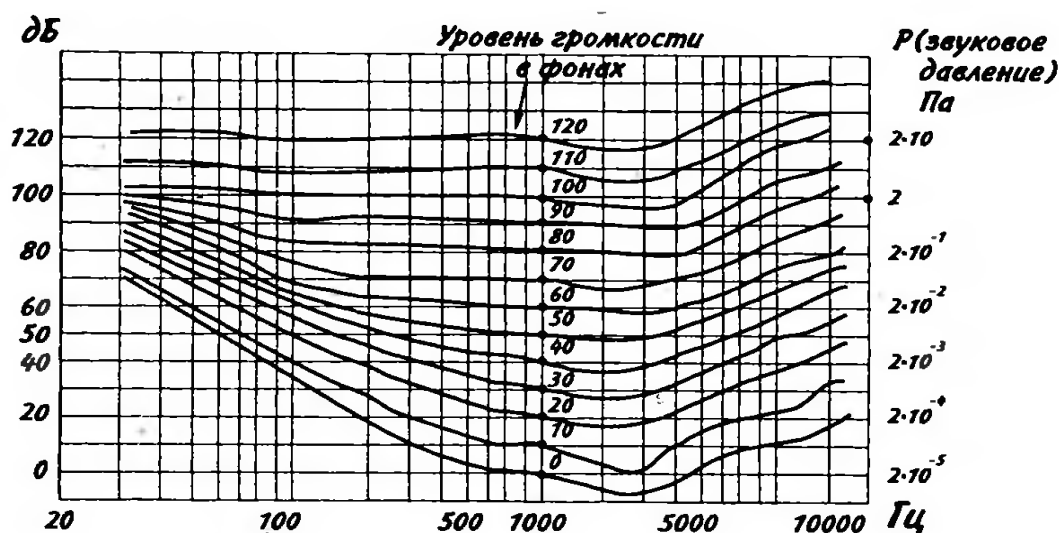


Рис. 1.2 Кривые равной громкости

Уровень громкости измеряется и настраивается в студиях при помощи

специального электроакустического прибора — *шумомера*, так как описанный метод субъективного сравнения измеряемого звука с эталонным тоном довольно трудоемок, не всегда точен и имеет в основном лишь теоретическое значение. При использовании шумомера измеряемый звук преобразуется микрофоном в электрические колебания. Затем, после усиления специальным усилителем, эти колебания измеряются прибором, отградуированным в единицах измерения уровня громкости. Приводим примерные уровни громкости некоторых типичных звуковых источников (табл. 2)

Таблица 2.

<i>Источник звука</i>	<i>Уровень громкости, фон</i>
Авиамотор на расстоянии 5 м	120
Фортиссимо оркестра	95-100
Шум движущегося поезда метро	90
Громкая радиомузыка, меццо-форте музыкального исполнения, шум на улице с движущимся транспортом	70-80
Аплодисменты	60-70
Пианиссимо музыкального исполнения, разговорная речь в жилой комнате	40-50
Тиканье часов на расстоянии 0,5 м	30

Слуховой аппарат

В настоящее время достаточно хорошо изучены области наружного, среднего и внутреннего уха, доводящие звуковые колебания воздушной среды до чувствительных окончаний слуховых волокон нервной системы. В значительно меньшей степени исследованы процессы в самой нервной системе. Известно, что здесь акустический сигнал преобразуется в электрический и в результате сложного взаимодействия в сфере высшей нервной деятельности создается звуковой образ, адекватный реальному.

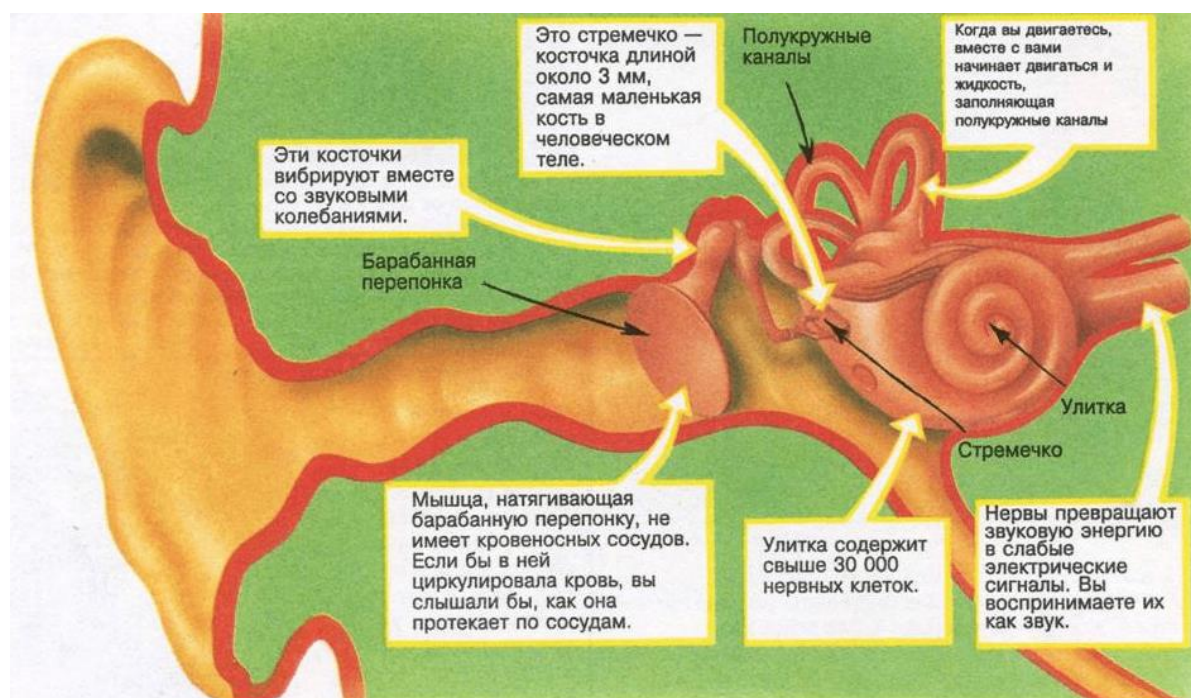


Рис. 1.3. Строение периферического отдела органа слуха

Ушная раковина направляет принимаемые акустические колебания в слуховой проход, заканчивающийся барабанной перепонкой. В слуховом проходе как в звуковом резонаторе, настроенном на частоты, близкие к 3 кГц, происходит примерно трехкратное усиление звукового давления, действующего далее на барабанную перепонку. Она образует границу с областью среднего уха и соединена здесь с костно-мышечным рычажным механизмом в виде молоточка и наковаленки. Мышечная ткань ножки наковаленки опирается на входной элемент внутреннего уха — мембрану овального окна внутреннего уха. Рычажная система молоточек-наковаленка выполняет роль трансформатора колебаний барабанной перепонки, повышая звуковое давление на мембране овального окна для наибольшей отдачи энергии из воздушной среды среднего уха, сообщаемой с

внешней средой через носоглотку, в область внутреннего уха, заполненную несжимаемой жидкостью — лимфой.

Структура внутреннего уха очень сложна и представляет собой сужающуюся к вершине трубку, свернутую в 2,5 витка в виде улитки, к которой примыкают каналы вестибулярного аппарата в виде трех колец.

Колебания через систему косточек (молоточек, наковальня, стремечко) передаются в главный отдел уха – улитку и там преобразуются в серии нервных импульсов, которые по нервным волокнам поступают в слуховой нерв, а по нему – в слуховые отделы головного мозга. Та часть внутреннего уха, где происходит преобразование механических колебаний в нервные импульсы, называется органом Корти. Результатом этого сложного процесса является преобразование входного сигнала в электрическую форму, и после этого с помощью слуховых нервов выполняется его передача к слуховым областям мозга, где и формируется окончательная реакция на звуковое воздействие

Пространственное восприятие звуков

Человек обладает способностью определять направление приходящих звуков. Органы слуха, воспринимая звуковые колебания от источников звука, расположенных справа и слева от слушателя, различают их и по амплитуде, и по фазе. При этом фазовые сдвиги приходящих звуковых волн зависят от расстояний между источником звука и левым и правым ухом. Различие по амплитуде звуковых давлений обусловлено тем, что голова оказывает определенное экранирующее действие, особенно заметное для звуковых волн, соизмеримых по длине волны с ее размерами (т.е. на высших частотах звукового диапазона). Оба этих явления воспринимаются человеком и служат информацией о месте источника звука в пространстве.

Способность человека определять место расположения (локализацию) источника звука, т.е. ощущать направление прихода звуковых волн, называется **бинауральным эффектом**, благодаря чему становится возможным пространственное слуховое ощущение. Определение «бинауральный» происходит от латинских слов *bini* — два и *auris* — ухо; соответственно, «моноуральный» означает «одним ухом».

Если источник сигнала расположить в свободном от отражений поле, то его звуковые волны будут поступать в оба уха с определенным различием во времени, по

интенсивности и спектру.

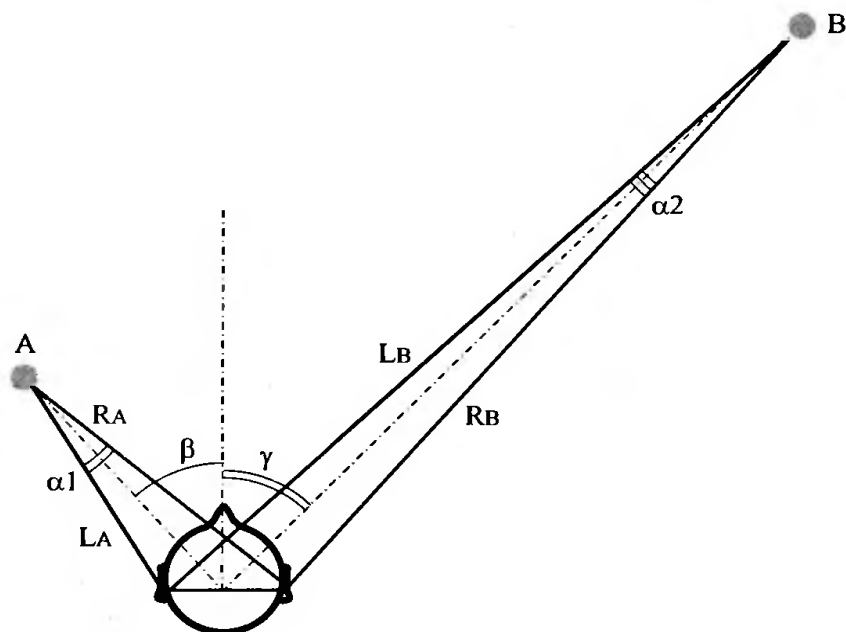


Рис. 1.4. Модель оценки дальности источника звука на основании анализа характеристик звуковых лучей

В итоге нервная система начнет анализировать пару подобных, но не идентичных звуков, которые сольются в единый образ. Первая фаза анализа носит название «бинауральное слияние» и определяет общие для обоих каналов характеристики с целью извлечения из сигнала полезной смысловой информации. Во второй фазе, наоборот, анализируются различия в каналах и решается задача получения информации о месторасположении источника звука и характеристиках окружающего пространства.

Наиболее важными для бинаурального слияния являются звуки с частотой ниже 1500 Гц. Можно провести характерный эксперимент: подадим в наушники два высокочастотных звука с разными частотами — они будут восприниматься на слух как отдельные звуковые сигналы. Если изменить условия опыта — промодулировать исходные сигналы низкочастотной огибающей, то оба сигнала сольются в единый слуховой образ. Результат свидетельствует о том, что нервная система для целей бинаурального слияния использует низкочастотную огибающую комплексного звука, то есть его макроструктуру, несмотря на то что детали составляющих комплексного звука — его микроструктура — различаются. Говоря научным языком, механизм бинаурального слияния звуков можно описать в виде математической модели, которая

основывается на поиске центральной нервной системой перекрестных корреляций между звуковыми сигналами в обоих ушах. Другими словами, используется поиск общих признаков в обоих каналах. Результаты поиска позволяют выделить из шума периодические компоненты сигналов, что, в частном примере, и позволяет понять речь.

Одновременно с выделением общих признаков ведется поиск отличий в тех звуковых сигналах, которые несут полезную для слушателя информацию. На основании анализа этих отличий нервная система способна сделать вывод о месторасположении источника в пространстве относительно слушателя и характеристиках окружающего пространства. При этом человек в состоянии определить, как удаленность источника звука, так и направление, с которого приходит звуковая волна

Лекция 2. Преобразователи звука

Преобразование акустических колебаний в электрический сигнал

Звук, который мы слышим, представляет собой упругие колебания окружающей нас воздушной среды (если мы, конечно, не нырнули с крутого бережка в прохладные воды тихой морской лагуны — там, естественно, звуковые колебания будут передаваться уже с помощью водной среды).

Однако, современные звукозаписывающие аппараты не способны реагировать на колебания давления окружающей их природной среды и «понимают» только колебания какой-нибудь электрической величины — тока, например, или напряжения. Чтобы преобразовать звуковые колебания воздушной среды в понятную каждому звукозаписывающему аппарату форму — электрический сигнал, требуется специальное устройство — микрофон.

В общем случае, микрофон — это прибор, реализующий функцию

$$U = f(P) \text{ или } I = f(P),$$

где U — напряжение в вольтах (В),

I — ток в амперах (А),

P — давление в паскалях (Па).

Способов преобразования акустических колебаний в электрический сигнал и конструкций микрофонов существует достаточно много. Рассмотрим некоторые из них. Первым устройством, использующимся только в качестве микрофона, стал **угольный** микрофон. Такой микрофон состоял из металлического корпуса, мембраны, в центре которой был укреплен подвижный электрод, неподвижного электрода, угольного порошка, изготавливаемого из лучших сортов каменного угля (антрацита).

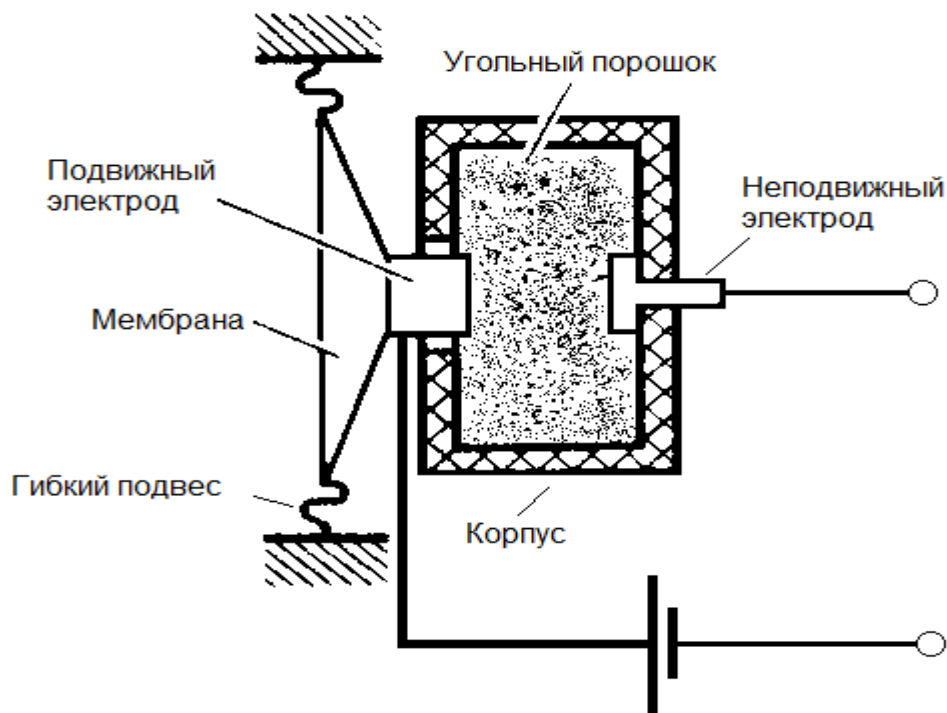
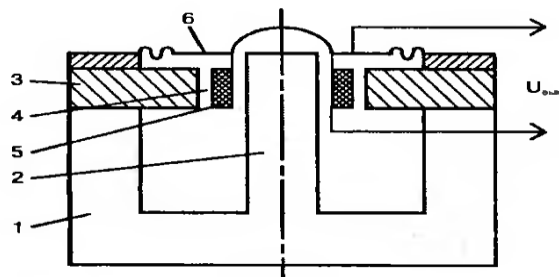


Рис 1. Конструкция угольного микрофона.

При воздействии звукового давления на подвижную мембрану, она прогибалась, уплотняя угольный порошок. Сопротивление его при этом падало и ток в цепи увеличивался. При уменьшении звукового давления подвижная мембрана распрямлялась, сопротивление угольного порошка увеличивалось и ток в цепи, соответственно, уменьшался. Таким образом, колебания тока в цепи с определенной степенью точности отображали колебания звукового давления и, индуцируя во вторичной обмотке согласующего трансформатора соответствующее напряжение, могли быть записаны на магнитофон или переданы в сеть связи.

Однако, угольные микрофоны никогда не отличались высокими техническими параметрами, поэтому к настоящему времени сохранились лишь в старых моделях телефонных аппаратов.

В профессиональной же звукозаписи наибольшее распространение получили электродинамические и конденсаторные микрофоны, которые обладают гораздо лучшими техническими характеристиками.



**Электродинамический
(катушечный) микрофон:**
 1 — постоянный магнит;
 2 — керн;
 3 — фланец;
 4 — кольцевой воздушный зазор;
 5 — звуковая катушка;
 6 — подвижная мембрана

Рис. 2. Устройство электродинамического (катушечного) микрофона

Устройство электродинамического (катушечного) микрофона показано на рис. 2. Оно напоминает устройство динамического громкоговорителя, поэтому последние часто используются в качестве микрофона — в переговорных устройствах, рациях — там, где качество сигнала не имеет решающего значения. Электродинамический микрофон имеет сильный постоянный магнит, напоминающий толстостенный стакан с круглым сердечником — керном в середине. К стороне, противоположной «дну» стакана прикреплен фланец — стальная накладка с круглым отверстием посередине. Между фланцем и керном магнита — узкий воздушный кольцевой зазор, в котором создается сильное магнитное поле. В этом кольцевом магнитном поле, не касаясь ни керна, ни фланца, находится звуковая катушка из изолированного провода. Катушка скреплена с подвижной мембраной, сделанной из алюминиевой фольги или специального пластика. Края мембраны гофрированы, благодаря чему она и скрепленная с ней звуковая катушка обладают подвижностью в осевом направлении.

Когда на мембрану воздействуют упругие колебания воздуха, вызванные акустическим сигналом, она начинает колебаться, увлекая за собой катушку. При этом провода катушки начинают пересекать силовые линии существующего в зазоре мощного магнитного поля, и в катушке индуцируется переменное напряжение той же частоты, что и у звуковых колебаний. Чем выше тон звука, тем выше частота колебаний напряжения. Чем громче звук, тем больше амплитуда электрических колебаний звуковой частоты.

Ленточный микрофон — разновидность электродинамического микрофона, в котором

используется тонкая мембрана (лента) из электропроводящего материала, помещенная между полюсами магнита. Ленточные микрофоны обычно являются двунаправленными. Звук хорошо снимается с источников, расположенных перед микрофоном и позади него, а сигнал источников, расположенных сбоку (под углом 90 градусов справа/слева), улавливаются слабо.

Устройство *конденсаторного микрофона* показано на рис. 3.

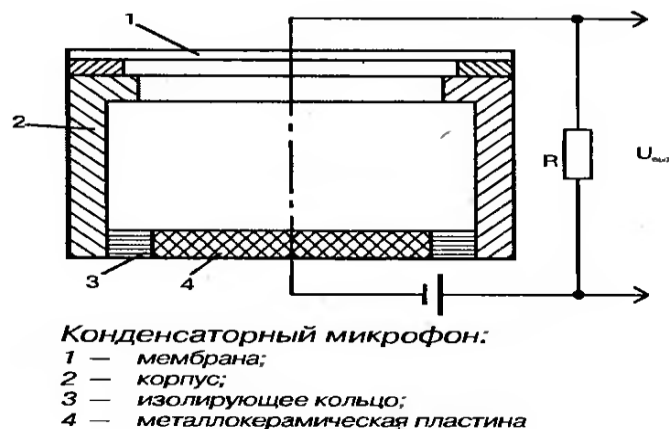


Рис. 3. Устройство конденсаторного микрофона

Основу его, как следует из названия, составляет конденсатор, одной из обкладок которого (неподвижной) является металлокерамическая пластина, другой — подвижная металлическая мембрана. Корпус микрофона изготовлен из керамики и отделен от металлокерамической пластины изолирующим кольцом. Последовательно с образованным таким образом конденсатором включен источник тока и нагрузочное сопротивление R . При воздействии на мембрану акустических колебаний она начинает изгибаться в ту или иную сторону, вызывая тем самым изменение емкости конденсатора. Заряд из конденсатора начинает перетекать в источник тока (при уменьшении емкости) и обратно (при увеличении емкости). Движения заряда есть не что иное как ток, который вызывает на включенном в цепь нагрузочном сопротивлении R переменное напряжение. Частота и амплитуда этого напряжения соответствует частоте и амплитуде воздействующего на микрофон акустического сигнала.

В последние годы широкое применение нашли *электретные микрофоны*, в которых используются некоторые диэлектрики, на поверхности которых после специальной обработки (нагрев и воздействие сильного электрического поля) появляются заряды. *Электретные микрофоны* по принципу работы являются теми же конденсаторными,

но постоянное напряжение в них обеспечивается зарядом электрета, тонким слоем нанесённого на мембрану и сохраняющим этот заряд продолжительное время (свыше 30 лет).

Кроме вышеперечисленных, встречаются еще **пьезоэлектрические микрофоны** — микрофоны, работающие на пьезоэлектрическом эффекте. При деформации пьезоэлектриков на их поверхности возникают электрические заряды, величина которых пропорциональна деформирующей силе. Пластины из искусственно выращенных кристаллов служат основным рабочим элементом пьезоэлектрических микрофонов. По характеристикам пьезоэлектрические микрофоны уступают большинству конденсаторных и электродинамических микрофонов, однако в некоторых сферах подобные микрофоны всё же применяются, например в качестве датчика гидрофоны, в бюджетных или устаревших гитарных звукоснимателях.

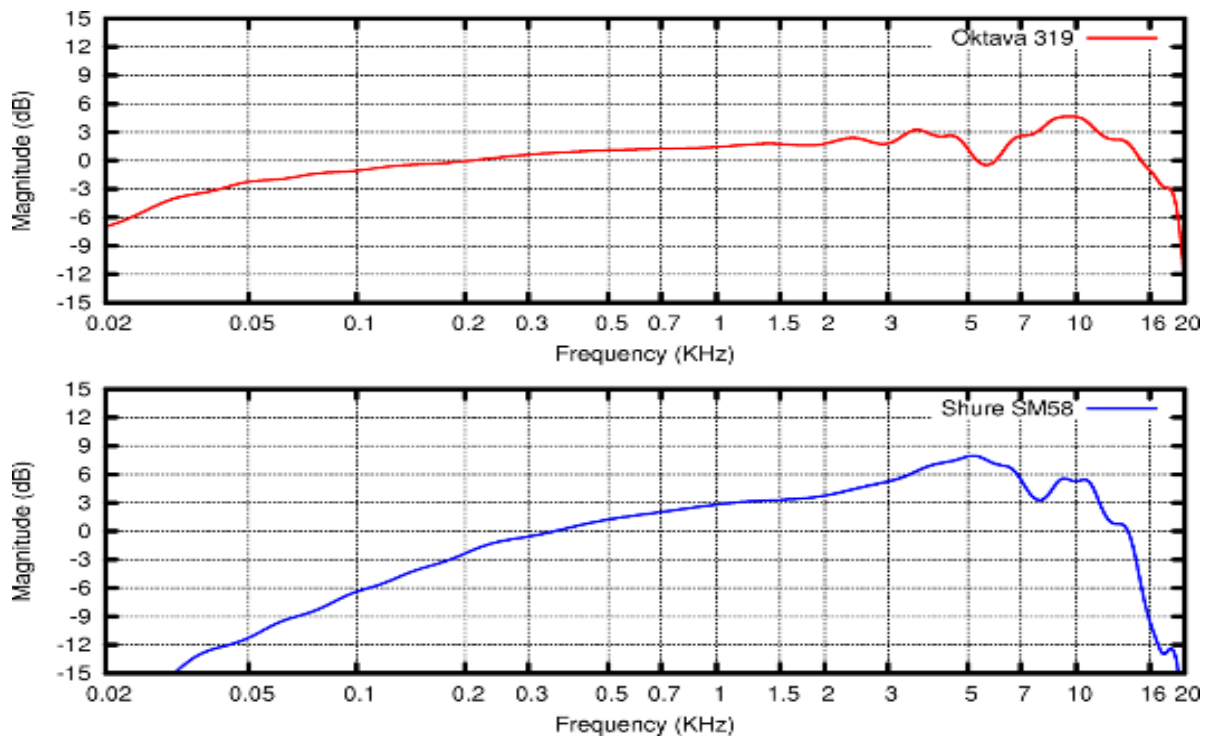
Технические параметры микрофонов

Чувствительность микрофона — отношение напряжения на выходе микрофона к воздействию на него звуковому давлению, выраженное в мВ/Па.

$$M_0 = U/P_0$$

Чувствительность современных микрофонов составляет от 1–2 (динамические микрофоны) до 10–15 (конденсаторные микрофоны) мВ/Па. Чем больше это значение, тем выше чувствительность микрофона. Таким образом, микрофон с чувствительностью –75 дБ менее чувствителен, чем –54 дБ, а с обозначением 2 мВ/Па менее чувствителен, чем 20 мВ/Па. Для ориентировки: –54 дБ это то же, что и 2,0 мВ/Па. Также надо учесть, что если у микрофона меньше чувствительность, это вовсе не означает, что он хуже.

Частотная характеристика (ЧХ) — это зависимость осевой чувствительности или ее уровня от частоты.



Неравномерность частотной характеристики — отношения чувствительности микрофона на определённой частоте к чувствительности на опорной частоте (в основном 1 кГц, измеряют в децибелах).

Диапазон (полоса) воспроизводимых частот - это диапазон звуковых частот, в котором неравномерность ЧХ относительно средней частоты 1000 Гц, как правило, не превышает 6 дБ.

Уровень собственных шумов микрофона — выраженное в децибелах отношение эффективного значения напряжения, обусловленного флуктуациями давления в окружающей среде и тепловыми шумами различных сопротивлений в электрической части микрофона, к напряжению, развиваемому микрофоном на нагрузке при воздействии на микрофон полного сигнала с эффективным давлением 1 Па.

Уровень звукового давления — показывает максимальный уровень громкости, который может принять микрофон. Измеряется в децибелах (Дб). Среднее значение – 100 Дб, повышенное значение – 130 Дб.

Характеристика направленности — зависимость чувствительности микрофона в свободном поле на определенной частоте от угла между осью микрофона и направлением на источник звука.

Диаграмма направленности — графическое изображение характеристики направленности в полярных координатах.

Типовые диаграммы направленности микрофонов показаны на рис. 4.

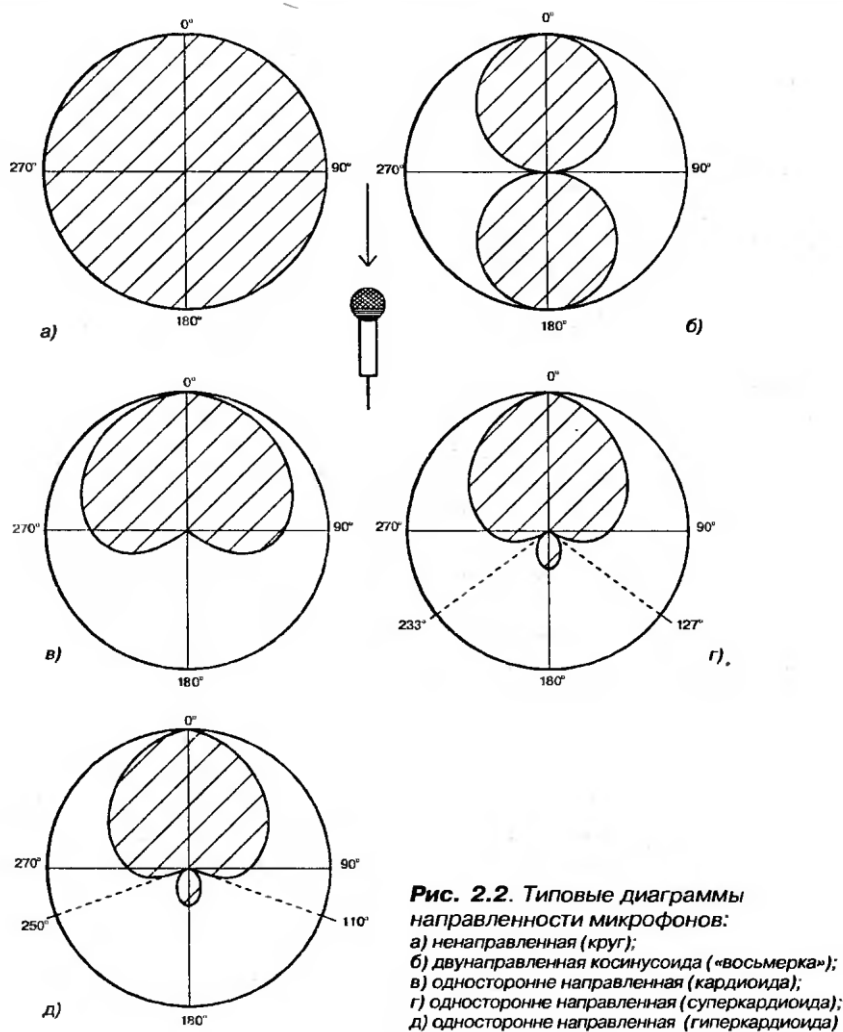


Рис. 4. Типовые диаграммы направленности микрофонов

Сравнительные характеристики основных типов микрофонов:

Тип микрофона	Диапазон воспринимаемых частот, <u>Гц</u>	Неравномерность частотной характеристики, <u>дБ</u>	Осевая чувствительность на частоте 1 000 Гц, <u>мВ/Па</u>
Угольный	300—3400	20	1000
Электродинамический катушечного типа	30—15 000	12	0,5 - 1,0
Электродинамический ленточного типа	60 —16 000	10	1 - 1,5
Конденсаторный	30—20 000	5	5
Пьезоэлектрический	100—5000	15	50

Конденсаторные микрофоны являются наиболее совершенными приборами, так как имеют широкий частотный диапазон, малую неравномерность частотной характеристики, низкий уровень нелинейных и переходных искажений, высокую чувствительность и низкий уровень собственных шумов. В этом их главное преимущество перед электродинамическими микрофонами, которые также имеют широкое распространение в технике звукозаписи. Номинальный диапазон частот у конденсаторных микрофонов может достигать величины 20...20000 Гц, неравномерность частотной характеристики чувствительности во всем диапазоне частот — не более 8 дБ.

Преобразование электрического сигнала в акустические колебания

Громкоговорители

Согласно принятому определению, под *громкоговорителями* понимаются пассивные электроакустические преобразователи, предназначенные для излучения звука в окружающее пространство. По способам излучения громкоговорители подразделяются на диффузорные — непосредственного излучения колеблющейся диафрагмой с гибкой подвеской и на рупорные — излучения с помощью жесткого рупора. Конструктивно каждый из громкоговорителей представляет собой совокупность двух независимых узлов — динамической головки и акустического оформления, согласованных по акустическим свойствам. Головка громкоговорителя — это собственно преобразователь сигналов звуковой частоты из электрической формы в акустическую — содержит все необходимые для этого конструктивные элементы, определяемые способом электроакустического преобразования. Акустическое оформление является элементом громкоговорителя, не участвующим в процессе указанного преобразования; оно лишь обеспечивает эффективное излучение звука в пространстве с помощью различного вида акустических экранов, ящиков, рупоров в разных вариантах построения.

Устройство громкоговорителей существенно зависит от вида главного элемента — головки, содержащей тот или иной электроакустический преобразователь. В настоящее время в качестве основы для построения громкоговорителей используют преимущественно электродинамические преобразователи (динамики). Соответственно и громкоговорители, укомплектованные головками с электродинамическими преобразователями, называются электродинамическими (сокращенно динамическими).

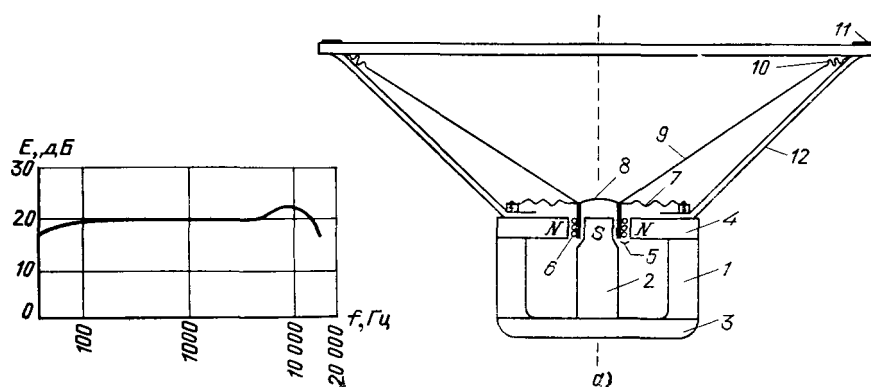


Рис. 5. Электродинамическая головка диффузорного громкоговорителя

Электродинамическая головка диффузорного громкоговорителя (рис. 4) состоит из магнитной системы в виде кольцеобразного постоянного магнита 1, обладающего большой коэрцитивной силой, круглого стального керна 2, проточенного в верхней части, нижнего 3 и верхнего 4 стальных фланцев. Верхний фланец имеет в центре отверстие, несколько большее диаметра керна, так что между фланцем и керном в верхней части образован воздушный зазор 5. В нем сконцентрировано радиальное постоянное магнитное поле. В зазоре располагается цилиндрическая катушка индуктивности 6 с небольшим числом витков, намотанная на каркасе, который скреплен с центрирующей гибкой шайбой 7 пылезащитным колпачком 8 и диффузором 9. Диффузор изготавливается из бумажной массы, обеспечивающей ему достаточную жесткость при малой массе. Обычно ему придают форму усеченного конуса, иногда с изогнутым (выпуклым) профилем. Диффузор по окружности крепится через гибкий «воротник» 10 к кольцу держателя 12. Подвижная система таким образом имеет лишь одну степень свободы движения — вдоль рабочей оси. От начала и конца катушки делаются гибкие выводы, соединяемые с выходом усилителя электрического сигнала.

Диффузор в таком громкоговорителе является одновременно и элементом колебательной механической системы, и излучателем звуковых колебаний в пространство. Колебания диффузора передаются частицам среды, прилегающим к поверхности диффузора, те в свою очередь воздействуют на смежные с ними частицы, и так в виде сжатий и разрежений образующиеся акустические волны распространяются от излучаемых участков поверхности преимущественно в нормальных к ним направлениях (частично воздействуя и на боковые частицы среды). Энергия колебательного процесса подвижной системы расходуется частично на преодоление механического сопротивления — трения воздуха в магнитном зазоре, ведущего к нагреву катушки, а частично на преодоление сопротивления колебаниям со стороны среды.

Входное сопротивление колебательной системы головки громкоговорителя можно представить в виде собственно электрического сопротивления катушки и сопротивления, вносимого элементами механической колебательной цепи, в данном случае включенными последовательно.

Рупорный громкоговоритель состоит из электродинамической головки прямого

излучения и рупора. Роль излучателя выполняет вибрирующая мембрана.

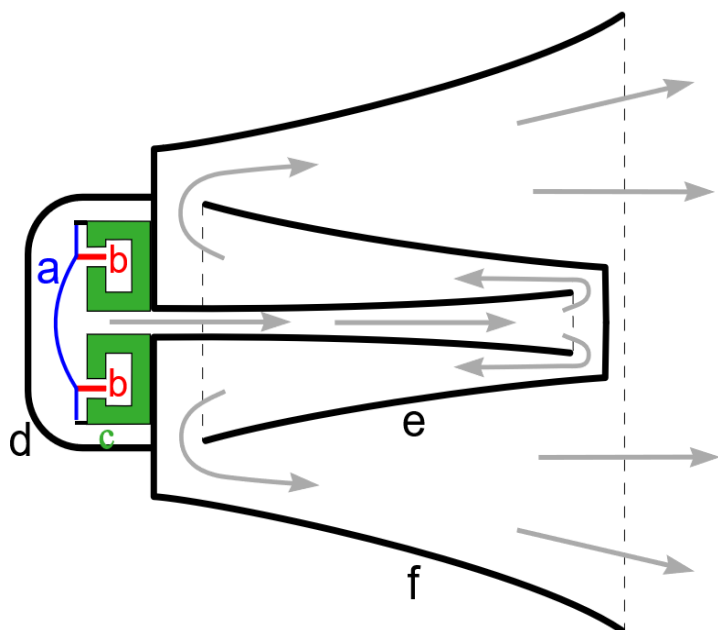
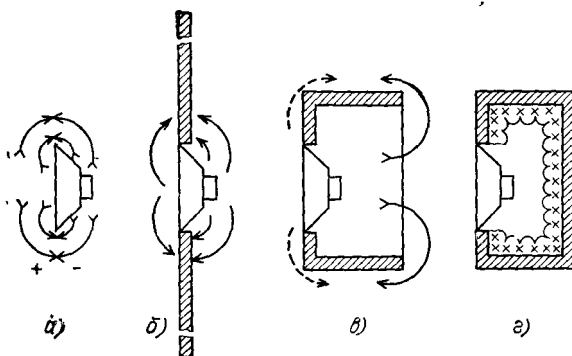


Рис. 6. Рупорный громкоговоритель

Рупорные громкоговорители чаще всего применяются в случаях, когда требуется большая громкость, но не требуется высокого качества звука — в таком случае достаточно просто создать рупорный громкоговоритель небольших габаритов, развивающий значительное звуковое давление при небольшой подводимой мощности (а значит — имеющий высокий КПД).

Главным недостатком диффузорных громкоговорителей является низкий коэффициент полезного действия — около 3...5%. Это объясняется недостаточной согласованностью между собой механического сопротивления подвижной системы и акустической нагрузки среды, а также недостаточно высокой эффективностью электродинамического преобразования.

Варианты акустического оформления головки громкоговорителя рассмотрены ниже:



a – в открытом пространстве

б - в отверстии плоского бесконечного акустического экрана

в - в корпусе с открытой задней стороной

г - в корпусе с закрытой задней стороной

Рис. 7 Варианты акустического оформления головки громкоговорителя

Громкоговоритель одностороннего действия (головка с экраном и головка в закрытом ящике) на низких частотах будет обладать ненаправленной характеристикой. Волны, излучаемые вдоль оси громкоговорителя, складываются в одинаковой фазе и взаимно усиливаются. Волны, излучаемые под некоторыми углами θ к рабочей оси, складываются с учетом различных фаз, так как они излучаются с разных участков поверхности диффузора и проходят пути, отличающиеся на разность хода крайних точек. Однако в области низких частот разность фаз получается очень малой и ею можно пренебречь. Для снижения интерференционных явлений и повышения эффективности использования энергии, излучаемой оборотной стороной диффузора, особенно в области нижних частот, иногда используют акустическое оформление в виде фазоинвертора — закрытого ящика с дополнительным отверстием на лицевой панели; к нему примыкает выходная часть акустического лабиринта, например, трубы, как на рис. 6.

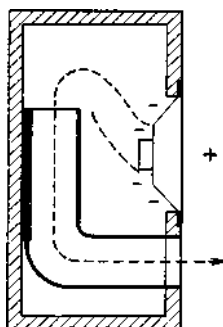


Рис. 8 Схема акустического фазоинвертора

Энергия звуковых колебаний, концентрирующаяся во внутреннем объеме ящика, через акустический лабиринт передается в сторону прямого излучения, и при надлежащей его длине фаза звуковых колебаний изменяется на обратную. Таким образом достигается возможность синфазного сложения прямого и обратного излучений между собой.

Акустические системы

Акустические системы (АС) представляют собой совокупность специально подобранных по частотным и другим свойствам электродинамических головок громкоговорителей диффузорного типа, заключенных в общий корпус с определенными акустическими параметрами. Каждая из головок охватывает определенную зону звукового диапазона; напряжение питания подается на головки (или их группы) через разделительные электрические фильтры, установленные в корпусе АС.

Чаще всего весь диапазон подразделяется на низко-, средне- и высокочастотный поддиапазоны. Соответственно для воспроизведения сигналов подбираются низко-, средне- и высокочастотная головки. Низкочастотная головка по сравнению с другими должна быть более мощной, поскольку чувствительность слухового аппарата в этой области понижена.

При комплектации акустических систем узкодиапазонными специализированными головками требуется согласование между полосой частот, которую они охватывают, и полосой частот электрических сигналов, которые следует подавать на каждую головку системы. В противном случае, если на вход каждой головки будет поступать более широкополосный сигнал, возможно появление существенных искажений. Для такого согласования применяют электрические пассивные фильтры, располагаемые обычно в одном корпусе с головками. Здесь же размещают регуляторы тембральной окраски звучания.

Для обеспечения малых нелинейных искажений номинальную мощность всех громкоговорителей выбирают с большим запасом, чтобы в реальной обстановке средняя мощность излучения была меньше указанной, а ее запас использовался при всплесках сигналов на уровне их максимальных значений.

Свойства громкоговорителей и АС принято оценивать многими параметрами и характеристиками. Из многочисленного перечня выделим лишь основные.

Частотная характеристика звукового давления (или просто частотная характеристика) — зависимость от частоты звукового давления, развиваемого громкоговорителем.

Номинальная мощность — максимальная мощность электрического сигнала, подводимого к громкоговорителю, ограничиваемая нелинейными искажениями, устанавливаемыми для данного громкоговорителя, при которой обеспечиваются его механическая и тепловая прочность.

Акустическая мощность — усредненная (во времени) мощность излучаемого громкоговорителем сигнала на частоте (или в полосе частот).

Полоса воспроизводимых частот — диапазон звуковых частот, в котором неравномерность частотной характеристики, как правило, не превышает 6 дБ. Колонки, как правило, имеют диапазон, ограниченный в области низких частот 40 — 60 Гц. Решить проблему воспроизведения низких частот позволяет использование сабвуфера.

Чувствительность звуковой колонки характеризуется звуковым давлением, которое она создает на расстоянии 1 м при подаче на ее вход электрического сигнала мощностью 1 Вт. В соответствии с требованиями стандартов чувствительность определяется как среднее звуковое давление в определенной полосе частот.

Чем выше значение этой характеристики, тем лучше АС передает динамический диапазон музыкальной программы. Разница между самыми «тихими» и самыми «громкими» звуками современных фонограмм 90 — 95 дБ и более. АС с высокой чувствительностью достаточно хорошо воспроизводят как тихие, так и громкие звуки.

Головные телефоны

Наушники (*тех. головные телефоны*) — устройство для персонального прослушивания звука. Наушники представляют собой пару небольших по размеру звукоизлучателей, надеваемых на голову или вставляемых прямо в ушные каналы. Наушники часто применяются в быту и в профессиональной деятельности для прослушивания музыки и речи, когда необходима мобильность или звукоизоляция от окружающего пространства. Наушники с прикреплённым к ним микрофоном образуют *головную гарнитуру*. Иногда в профессиональной деятельности вместо пары используется один отдельный наушник, он называется *монитор*. В технической литературе вместо слова «наушники» традиционно используется термин «головные телефоны». Термин возник, скорее всего, в результате буквального прочтения английского слова *headphones* (англ. *head* — голова, *phone* — телефон). Классификация наушников.

По конструкции излучателя

- *динамические* — используют электродинамический принцип преобразования. Представляет собой миниатюрный электродинамический диффузорный громкоговоритель, заключенный в закрытый (иногда с отверстиями) корпус, прилегающий через гибкую прокладку к уху. В такой конструкции хорошо выявляются положительные свойства, присущие диффузорным электродинамическим головкам. Диапазон частот, охватываемый типовыми головными телефонами, достигает 20 ... 20 000 Гц при малых отклонениях от горизонтальной линии частотной характеристики. При небольшой мощности головные телефоны обеспечивают достаточное звуковое давление для создания у слушателя ощущения высокой громкости в очень широком динамического диапазона без нелинейных искажений. Самый распространённый тип наушников.
- *электромагнитные с уравновешенным якорем* — представляет собой миниатюрный электромагнитный громкоговоритель в котором звуковые колебания создаются за счёт движения мембраны из магнитного материала либо металлического якоря в поле неподвижного электромагнита. Основной деталью

является П-образный якорь из ферромагнитного сплава. В разговорной речи такие наушники часто называют «арматурными» из-за созвучия английского слова *armature* (здесь «якорь»);

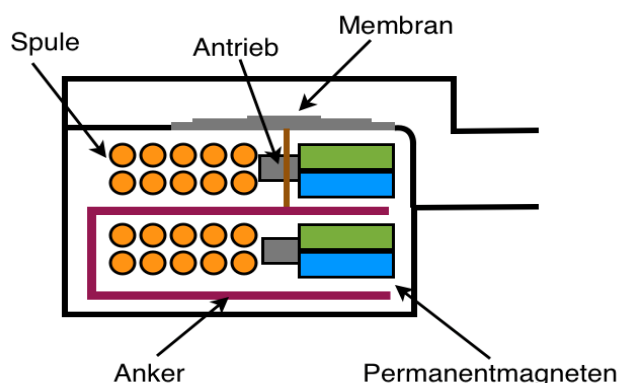


Рис. 9 Конструкция излучателя с уравновешенным якорем

По типу конструкции (виду)

- *внутриканальные* (обиходное название — «затычки», «капельки», «вакуумки», «бочки», «беруши») — вставляются в ушной канал; по типу преобразования это, как правило, электромагнитные с уравновешенным якорем
- *вставные* (обиходное название — «вкладыши», «пуговки») — вставляются в ушную раковину;
- *накладные* — накладываются на голову, полностью закрывая ухо; по типу преобразования это, как правило, *динамические*.

Таковые, в свою очередь, делятся на наушники:

- *открытого типа* — наушники пропускают внешние звуки через специальные широкие отверстия. Многие слушатели отмечают звук открытых наушников как более естественный, с широкой звуковой сценой, и менее утомительный по сравнению со звуком закрытых наушников. Однако, в сравнении с наушниками закрытого типа, открытые наушники не обеспечивают хорошей звукоизоляции и, как правило, слабее воспроизводят низкие частоты;

- *закрытого типа* — чаши наушников не имеют внешних отверстий. За счёт этого наушники не пропускают внешние шумы и обеспечивают максимальную звукоизоляцию, что позволяет использовать их в шумных условиях, а также в тех случаях, когда необходимо полностью сосредоточиться на прослушивании или

прослушать конфиденциальную информацию. При высоком качестве звука, основные недостатки закрытых наушников заключаются в высокой утомляемости слуха и дискомфорте для головы в силу отсутствия циркуляции воздуха из-за плотного прилегания амбушюров (при плохом же прилегании амбушюров ухудшается воспроизведение низких частот) и, как правило, большего давления на голову, чем у открытых наушников;

- *полуоткрытого типа* (или полужакрытого типа) — открытые наушники, но лишь с небольшими по размеру внешними отверстиями, призванные объединить в себе достоинства наушников закрытого и открытого типа.

Лекция 3. Аналого-цифровое преобразование.

Все физические процессы, происходящие в окружающем нас мире, являются непрерывными, т.е. изменение во времени любой физической величины от одного значения к другому происходит постепенно и включает в себя несчетное множество всех промежуточных значений этой величины. Отображение такого процесса путем аналогичного изменения какой-нибудь другой физической величины, например, электрического тока или напряжения, тоже должно происходить непрерывно и включать в себя все промежуточные значения отображаемого процесса.

Несчетное множество значений какой-либо физической величины, которые она принимает с течением времени, называется континуальным (непрерывным) множеством, а электрический сигнал, отображающий изменение этой величины, называется континуальным или аналоговым сигналом.

Звук, который мы слышим и который представляет собой изменения давления в воздушной среде, является непрерывным или аналоговым процессом. Электрический сигнал, отображающий такой процесс, тоже будет иметь непрерывный характер. Это типичный аналоговый сигнал.

Если требуется сохранить звуковую программу для прослушивания ее в будущем, то эту программу тем или иным способом записывают на какой-нибудь носитель информации. Если этот носитель предполагает аналоговый способ записи (т.е. путем изменения одного из параметров этого носителя) с сохранением всего множества значений исходного сигнала, то теоретически при воспроизведении следовало бы ожидать абсолютно точного соответствия между записанным и воспроизведенным сигналами. Но это только теоретически. В действительности, каждый носитель обладает свойством вносить свои изменения в сделанную на нем запись. Такие изменения возникают как под воздействием внешних факторов (влияния окружающей среды или контакта с воспроизводящим устройством), так и в силу физических свойств самого носителя. Поскольку запись аналоговая, т.е. содержит в себе бесчисленное множество значений исходного сигнала, то малейшие изменения характеристик носителя приводят к пагубным и необратимым последствиям — в записанном

материале появляются искажения, т.е. воспроизведенная звуковая программа с течением времени будет все больше и больше отличаться от ее оригинального звучания.

Справиться со всеми вышеперечисленными недостатками помогает преобразование аналогового звукового сигнала в цифровую форму, т.е. в виде последовательности двоичных кодов.

Преобразование звукового сигнала в цифровую форму заключается в измерении мгновенных значений его амплитуды через равные промежутки времени и представлении полученных значений, называемых отсчетами, в виде последовательности чисел. Такая процедура называется **аналого-цифровым преобразованием**, а устройство для ее реализации — **аналого-цифровым преобразователем (АЦП)**. Числа, полученные в результате аналого-цифрового преобразования, выражаются в двоичной системе исчисления, т.е. в виде комбинации всего двух цифр — нулей (0) и единиц (1).

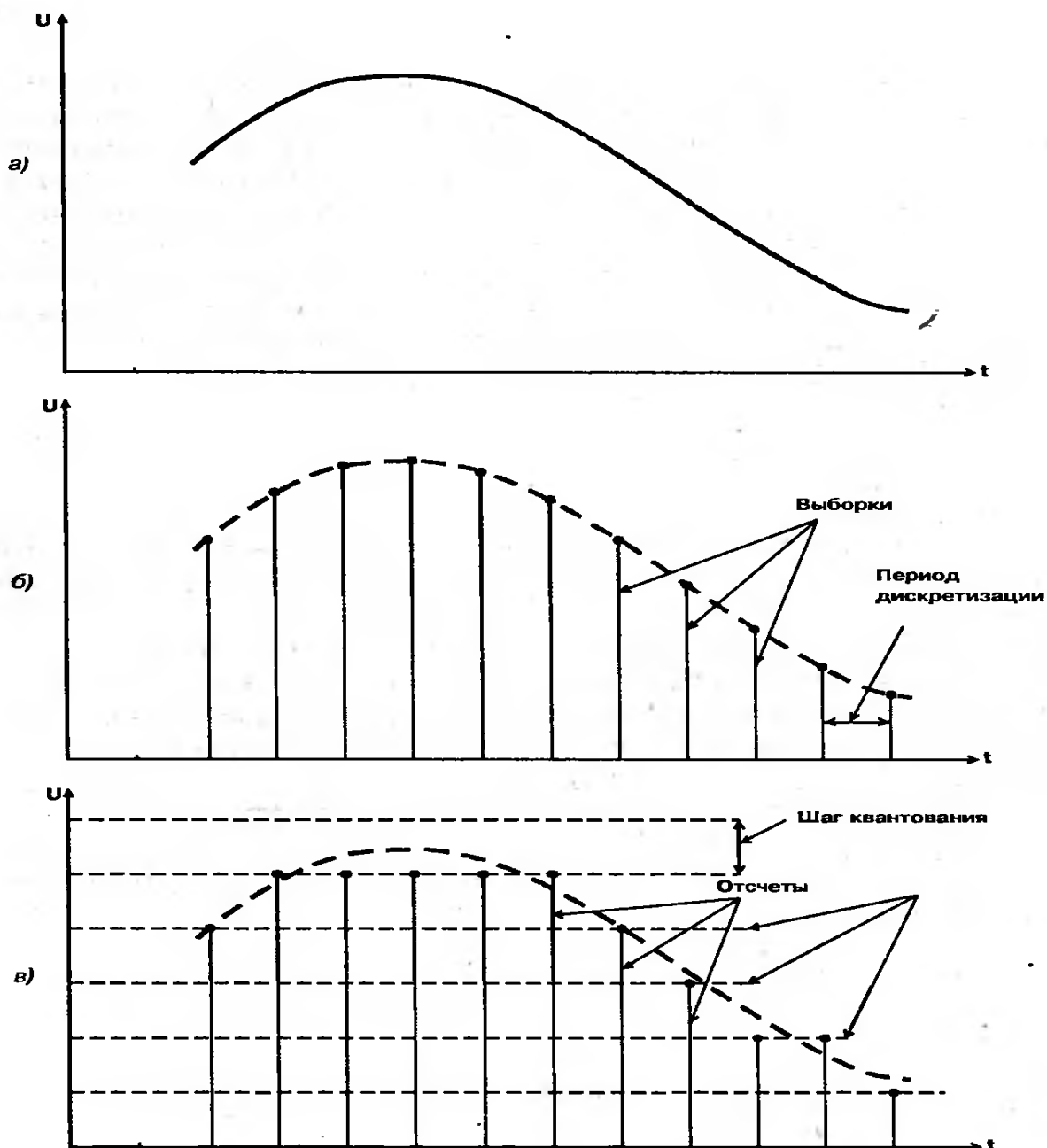


Рис. 1 Аналого-цифровое преобразование:

- а — исходный аналоговый сигнал;
 б — дискретизация;
 в — квантование

Процесс преобразования непрерывного аналогового сигнала в последовательность его мгновенных значений (выборок) называется **дискретизацией** (рис. 1, б). Определение численного значения величины выборки (отсчета) называется **квантованием**. Для этого весь диапазон возможных изменений амплитуды преобразуемого сигнала делится на множество уровней квантования, количество которых определяется разрядностью используемого при этом двоичного числа. Чем больше число разрядов квантования, тем

меньше расстояние между уровнями квантования (шаг квантования) и тем выше получается точность преобразования. В процессе квантования за величину выборки (отсчет) принимается номер ближайшего уровня квантования (рис. 1.в).

Скорость следования отсчетов в секунду называется частотой дискретизации, а время между двумя соседними отсчетами — *периодом дискретизации*.

Дискретизация по времени

Выбор частоты дискретизации в общем случае определяется известной теоремой Котельникова (теоремой отсчетов), которая в оригинале звучит так:

«Если наивысшая частота в спектре функции $S(t)$ меньше чем fm , то функция $S(t)$ полностью определяется последовательностью своих значений в моменты, отстоящие друг от друга не более чем на $1/(2fm)$ секунд».

В рассматриваемом случае под функцией $S(t)$ следует понимать непрерывный аналоговый звуковой сигнал, а под частотой fm — наивысшую частоту требуемого звукового диапазона. Если необходимо точно отобразить аналоговый сигнал в диапазоне до fm , то отсчеты должны следовать с периодом, по крайней мере, в два раза меньшим, чем период частоты fm . Иными словами, частоту дискретизации следует выбирать так, чтобы она была, по меньшей мере, в два раза выше максимальной частоты звукового диапазона.

При этом минимально возможная частота дискретизации называется частотой Найквиста

$$F_n = 2fm.$$

В зарубежной литературе в аналогичных случаях ссылаются на теорему Шеннона, которая, по сути, имеет тот же смысл, что и теорема Котельникова.

Дальнейшие рассуждения будут более понятными, если взглянуть на полную схему тракта аналого-цифрового преобразования, показанную на рис. 2.

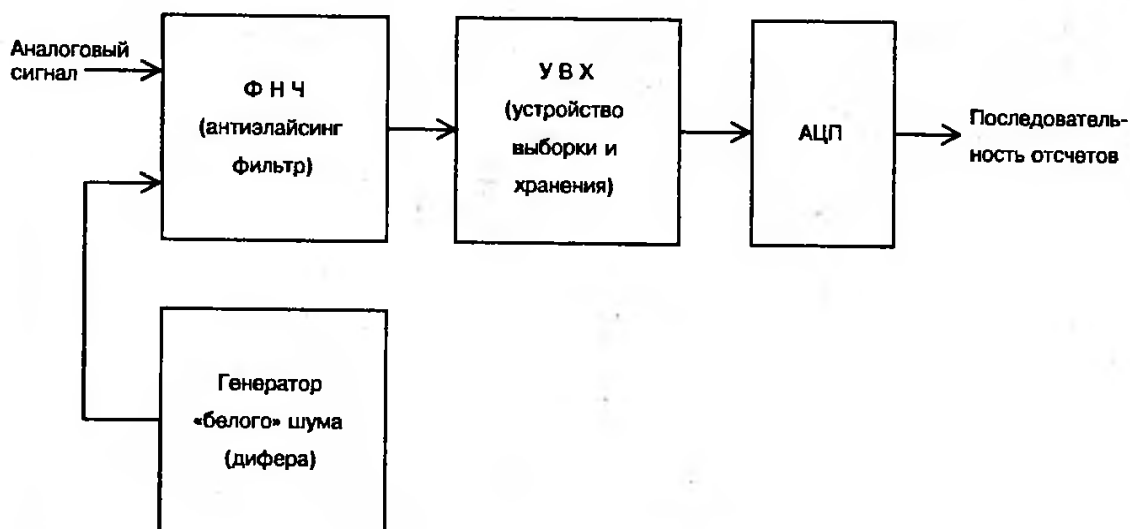


Рис. 2 Структурная схема АЦП

Но для начала обратим внимание на то, что процесс аналого-цифрового преобразования требует времени, а преобразуемый сигнал непрерывно изменяется, и если не принять никаких мер для фиксации его значений на время преобразования, то результат будет содержать ошибку. Чтобы такого не происходило, на входе АЦП размещают так называемое устройство выборки и хранения (УВХ), которое в общем случае представляет собой аналоговое запоминающее устройство (см. рис.3). В качестве элемента памяти в нем используется высококачественный конденсатор.

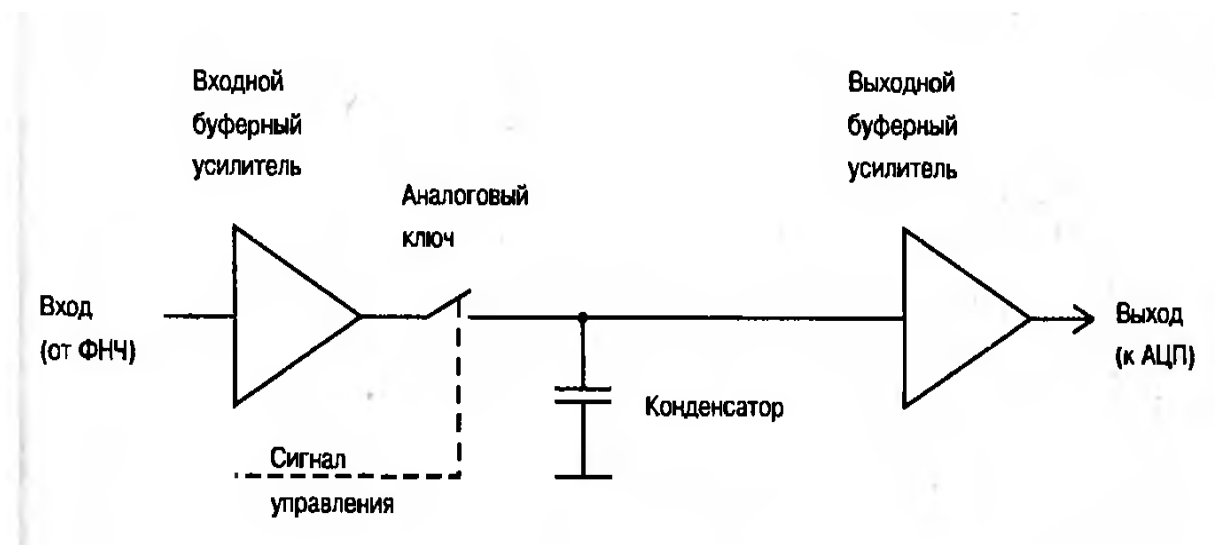


Рис. 3 Схема УВХ

Когда ключ замыкается (время выборки T_v) напряжение на конденсаторе изменяется в соответствии с изменением входного сигнала. При размыкании ключа (время хранения T_x) напряжение на конденсаторе фиксируется и сохраняется без изменения на все время, пока АЦП осуществляет преобразование. Точность УВХ практически определяет точность преобразования и должна быть не хуже точности, соответствующей разрядности применяемого АЦП.

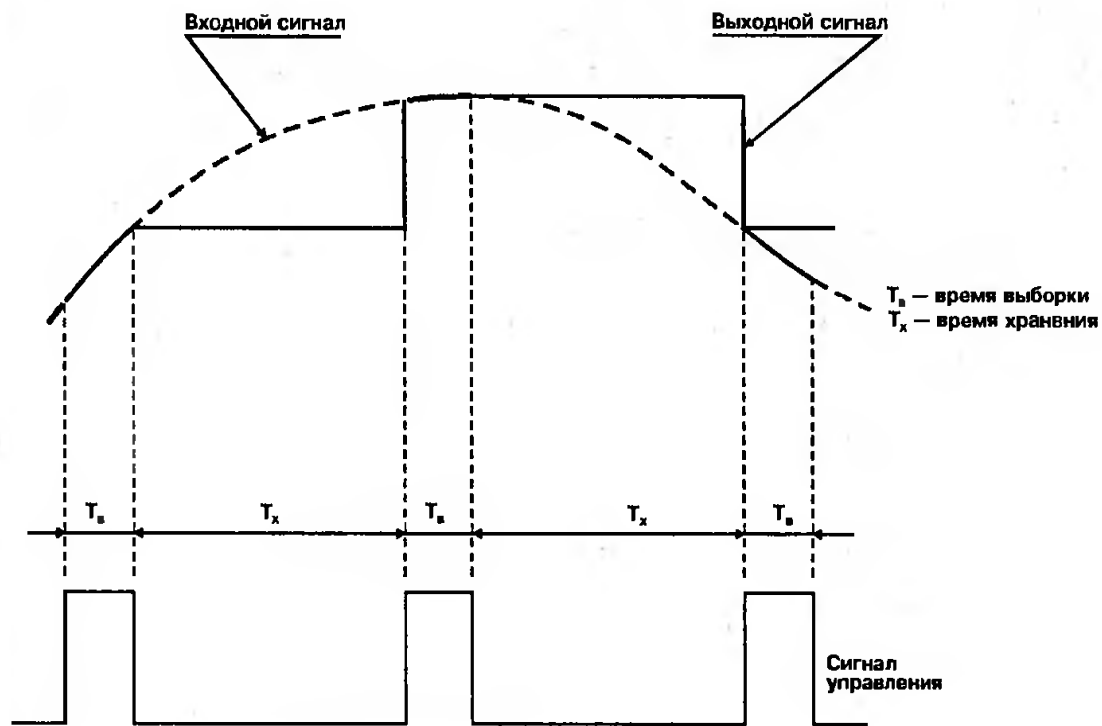


Рис. 4 Временная диаграмма работы УВХ

На практике минимальная частота дискретизации F_d :

$$F_d = (2,2...2,4)f_m.$$

Как правило требуемая полоса звуковых частот ограничивается 20кГц, а минимальная частота дискретизации при этом выбирается равной 44,1 или 48 кГц.

Это обусловлено тем, что между наивысшей частотой звукового диапазона f_m и половиной частоты дискретизации $F_d / 2$ должен быть некоторый интервал, в который нужно поместить срез амплитудно-частотной характеристики (АЧХ) фильтра низких частот (ФНЧ), расположенного на входе блока аналого-цифрового преобразования. Этот ФНЧ, который называется *антиэлайсинг* фильтром, нужен для того, чтобы ни одна составляющая спектра выше $F_d/2$ не попала на преобразователь. Дело в том, что спектр дискретизованного сигнала обладает периодической структурой.

Кроме низкочастотной части, отображающей сам звуковой сигнал, он имеет еще и высокочастотные компоненты в виде боковых полос с центрами в точках, кратных частоте дискретизации (рис. 5, а).

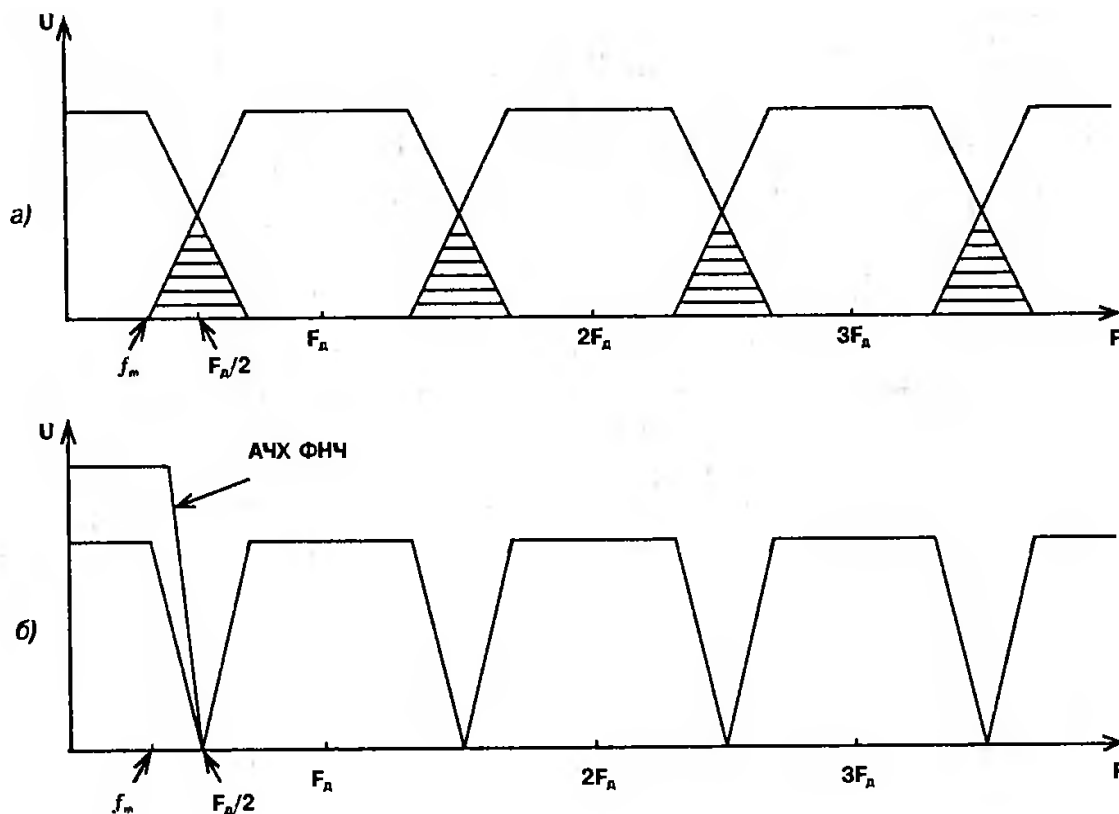
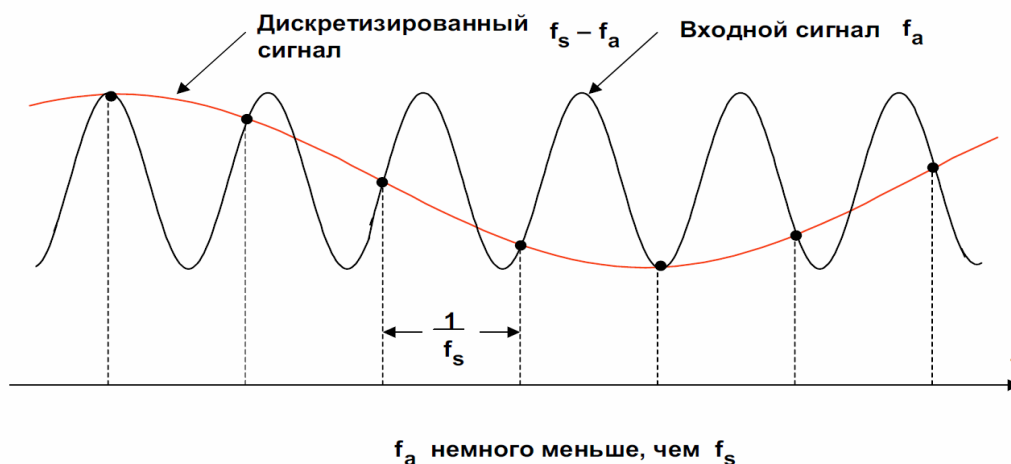


Рис. 5. Спектр дискретизованного сигнала: а — без ФНЧ; б — с ФНЧ

Если спектр звукового сигнала перед преобразованием не ограничить, то его высокочастотная часть может наложиться на смежную боковую полосу. При этом в преобразованном сигнале возникнут неустраняемые искажения *субдискретизации* в виде паразитных высокочастотных составляющих (рис. 5, а). Звучание фонограммы будет безнадежно испорчено.

Эффект наложения спектров во временной области



Квантование по уровню

При квантовании по уровню вырабатываются двоичные числа, которые представляют значения отсчётов аналогового сигнала. Двоичные числа являются цифровым представлением напряжения аналогового звукового сигнала в моменты дискретизации по времени. Количество битов N , использованные для кодирования отсчётов, звукового сигнала, называется разрядностью квантования по уровню. Аналогично тому, как частота дискретизации определяет ширину полосы частот цифрового тракта, разрядность квантования по уровню определяет её динамический диапазон, разрешающую способность и уровень нелинейных искажений.

Поскольку в процессе квантования отсчеты могут принимать только значения кратные шагу квантования Δ , то при оценке истинного значения выборки неизбежно будет возникать некоторая ошибка q (рис. 6).

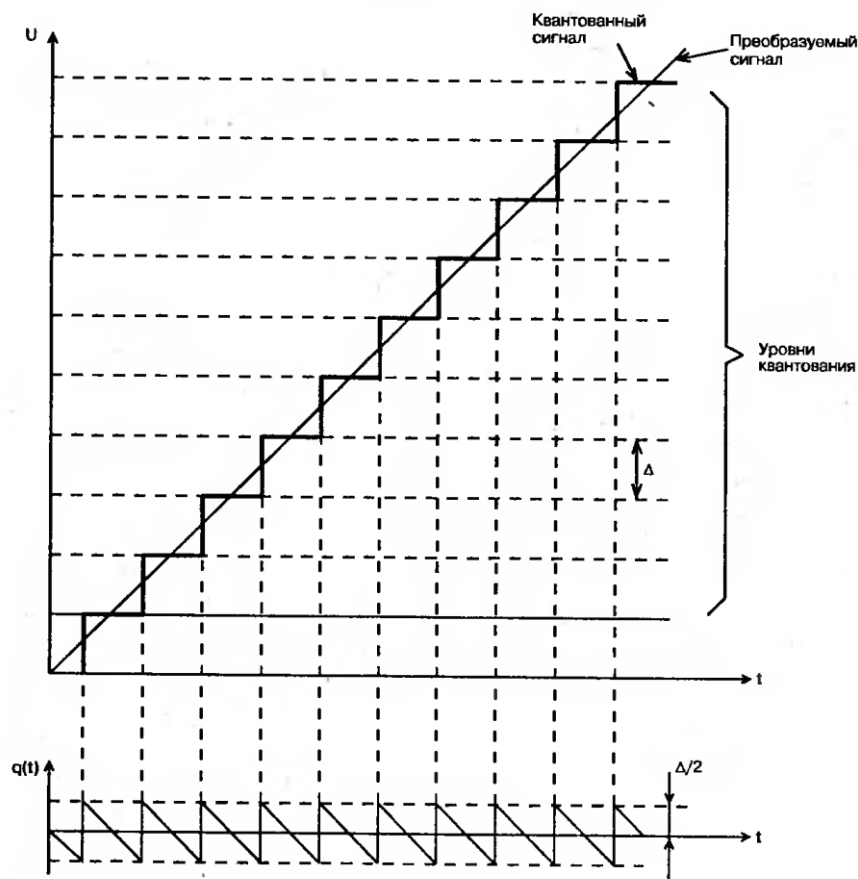


Рис. 6. Шум квантования

Очевидно, что величина ошибки равна половине шага квантования и не зависит от

уровня квантуемого сигнала. Функцию $q(t)$ принято называть шумом квантования. Шум квантования будет тем ниже, чем меньше шаг квантования или, что то же самое, чем больше число разрядов квантования. Относительная величина максимальной погрешности квантования равна

$$\delta = \frac{1}{2^{N-1}}$$

где N — число разрядов квантования. Этой же величиной, представленной в логарифмических единицах (децибелах), оценивается уровень шумов квантования АЦП.

Очевидно также, что влияние шума сильно зависит от уровня преобразуемого аналогового сигнала. Если его амплитуда мала, то возникают характерные искажения, обусловленные появлением высших гармоник из-за зубчатой формы шума квантования. При этом на слух это воспринимается именно как искажения, а не как шум.

Ослабить влияние таких искажений можно, как ни парадоксально это звучит, с помощью добавления другого шума. Если подмешать во входной сигнал так называемый «белый» шум (шум, значения которого некоррелированы, т. е. амплитуда которого практически постоянна в широком диапазоне частот), то корреляция (связь) между шумами квантования и амплитудой сигнала нарушается. При этом воспроизведенный сигнал уже не будет выглядеть искаженным. Добавление такого шумоподобного маскирующего сигнала (*дифера*) является важной частью процесса преобразования. Дифер так же повышает разрешающую способность системы по отношению к сигналам низкого уровня.

С числом разрядов квантования N физически связан динамический диапазон D звукового сигнала.

$$D = 6N + 1,76 \text{ дБ.}$$

Следовательно, для цифровых систем звукозаписи с 16-разрядным квантованием $D = 98$ дБ.

АЦП для цифровой обработки звука

АЦП последовательного приближения. Классическим методом аналого-цифрового преобразования является метод последовательного приближения.

Структурная схема АЦП, использующего такой метод, представлена на рис. 7, а временная диаграмма его работы — на рис. 8.

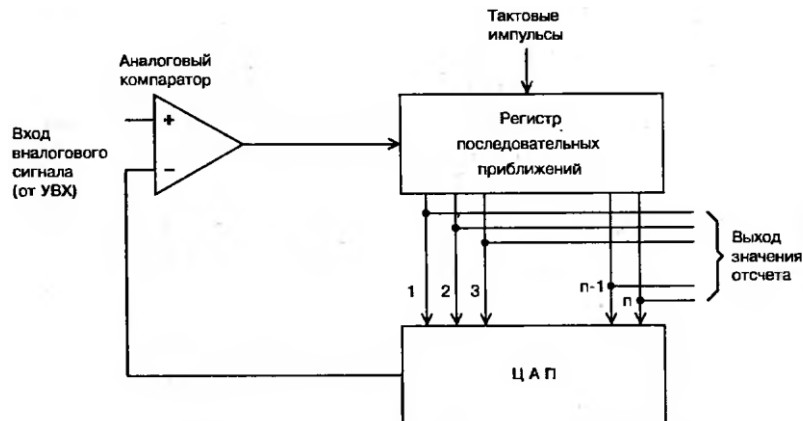


Рис. 7. Структурная схема АЦП последовательного приближения

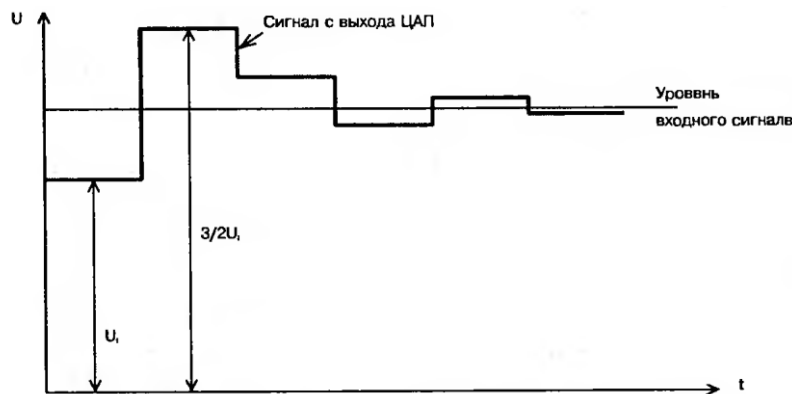


Рис. 8. Временная диаграмма АЦП последовательного приближения

Аналоговая выборка от УВХ поступает на первый вход аналогового компаратора. Первым тактовым импульсом в регистр последовательных приближений записывается код, который, попадая на входы ЦАП, вызывает появление на его выходе напряжения U . Это напряжение поступает на второй вход компаратора, который сравнивает оба напряжения. Если напряжение аналоговой выборки больше U , то в старший разряд регистра записывается 1, а если меньше — то 0. В зависимости от этого результата, после второго тактового импульса на выходе регистра последовательных приближений появится уже другой код, который вызовет изменение значения напряжения на выходе

ЦАП и оно станет равным некоторому значению U_2 . Если после первого такта в старший разряд была записана 1, то $U_2 = 3/2 U$, а если 0 — то $U_2 = 1/2 U$. Снова компаратор производит сравнение, а результат записывается в следующий разряд регистра. Этот процесс будет продолжаться с постепенным приближением напряжения на выходе ЦАП к значению аналоговой выборки столько раз, сколько разрядов имеет используемый АЦП.

Сигма-дельта АЦП. В последнее время все более широкое распространение получают методы А/Ц и Ц/А преобразования на основе использования сигма-дельта модуляции, когда квантование осуществляется всего одним разрядом, но с частотой в десятки и сотни раз превышающей частоту Найквиста (F_H). В процессе такого преобразования анализируется не амплитуда аналогового сигнала, а направление ее изменения. Если амплитуда возрастает, то результатом преобразования будет 1, а если уменьшается — то 0. Нулевой уровень кодируется чередующимися нулями и единицами.

Сигма-дельта (или дельта-сигма) модуляция известна достаточно давно — еще с 1962 года. Однако, практическое применение ее в системах высококачественного записи и воспроизведения звука до недавнего времени было затруднено тем, что не было достаточно быстродействующей элементной базы.

Основная идея такого метода состоит в том, что спектр шума квантования, возникающего в процессе дискретизации с низким разрешением, преобразуется так, что в полосе звуковых частот его уровень понижается, а в области высоких частот (за пределами основной полосы) повышается. Затем полученный цифровой поток обрабатывается прореживающим фильтром нижних частот (фильтр-дециматор) с получением в результате последовательности отсчетов необходимой разрядности, следующих с выбранной частотой дискретизации.

Избыточная дискретизация, цифровая фильтрация, формирование шума и прореживание

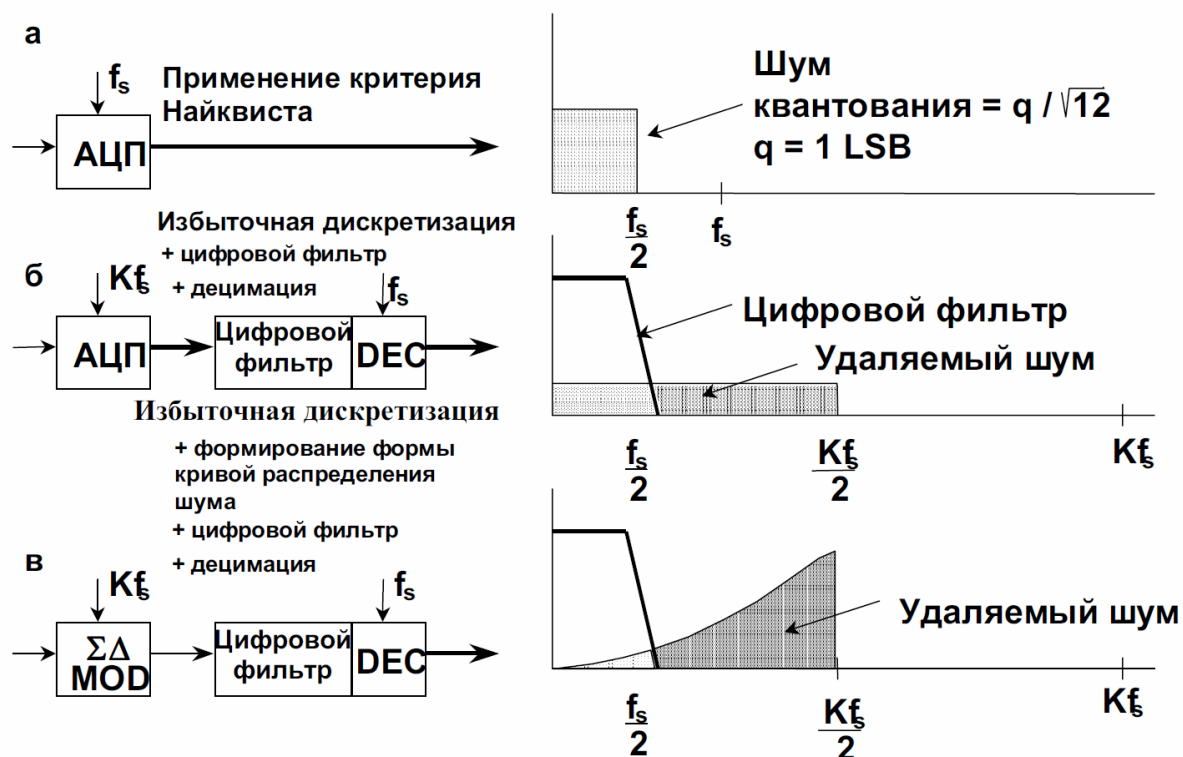


Схема блока аналого-цифрового преобразования с использованием сигма-дельта модуляции приведена на рис. 9.

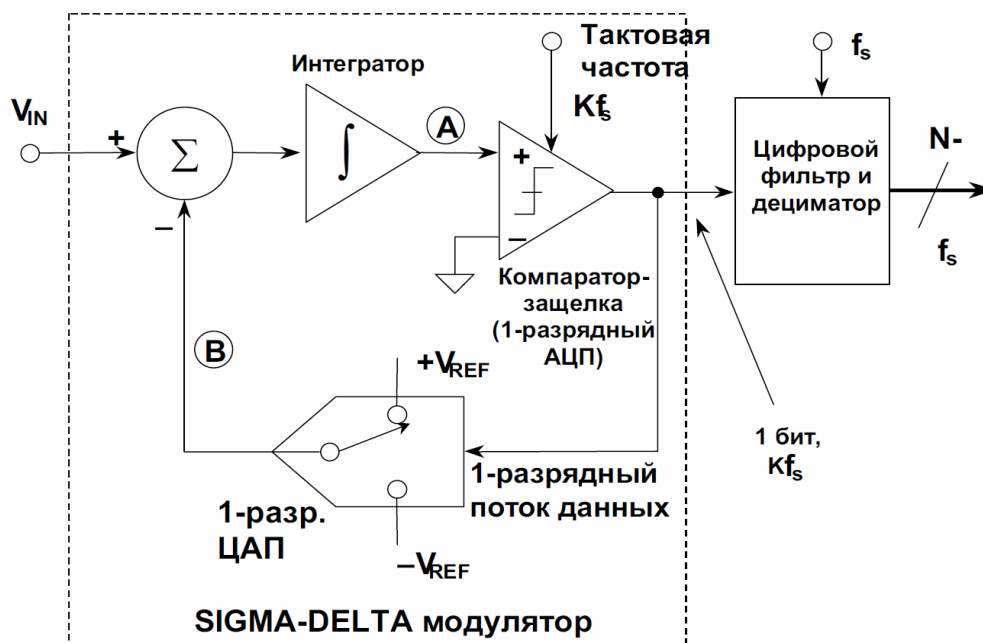


Рис. 9. Схема сигма-дельта АЦП

Не вдаваясь в детали, работу сигма-дельта АЦП можно описать следующим образом.

Представим, что постоянное напряжение подается на вход V_{in} . Сигнал на выходе интегратора в точке А при этом постоянно нарастает или убывает. С выхода компаратора сигнал подается обратно через одноразрядный ЦАП на вычитающий вход в точке В. Благодаря отрицательной обратной связи, соединяющей выход компаратора через одноразрядный ЦАП с точкой суммирования, среднее значение постоянного напряжения в точке стабилизируется на уровне V_{in} . Вследствие этого, среднее выходное напряжение ЦАП равняется входному напряжению V_{in} . В свою очередь, среднее выходное напряжение ЦАП определяется плотностью потока единиц в одноразрядном потоке данных, следующего с выхода компаратора. Когда значение входного сигнала увеличивается до $+V_{ref}$, число единиц в последовательном потоке данных увеличивается, а число нуле уменьшается. Точно так же, когда значение сигнала приближается к отрицательному значению $-V_{ref}$ число единиц в последовательном потоке данных уменьшается, а число нулей увеличивается. Попросту говоря, в последовательном потоке разрядов на выходе компаратора содержится среднее значение входного напряжения. Цифровой фильтр-дециматор обрабатывают последовательный поток битов и выдают окончательные выходные данные.

В отличии от многоразрядных АЦП, здесь для выделения полосы частот звукового сигнала не требуется применения сложного антиэлайсинг-фильтра. Вполне достаточно простого ФНЧ 3-го порядка. Кроме того, не требуется и применения устройства выборки и хранения (УВХ), так как преобразование аналогового сигнала осуществляется непосредственно, без предварительной фиксации величины выборки.

Лекция 4. Цифро-аналоговое преобразование

Процесс обратного преобразования последовательности цифровых отсчетов в аналоговый сигнал называется цифро-аналоговым преобразованием, а устройства для его осуществления — цифро-аналоговыми преобразователями (ЦАП).

На выходе ЦАП получается ступенчатый аналоговый сигнал, величина ступенек которого равна численному значению соответствующих отсчетов. Двоичные числа являются цифровым представлением напряжения аналогового звукового сигнала в моменты дискретизации по времени. Количество битов, использованные для кодирования отсчётов, звукового сигнала, называется разрядностью квантования по уровню. Разрядность квантования по уровню определяет её динамический диапазон, разрешающую способность и уровень нелинейных искажений ЦАП.

Если необходимо точно отобразить аналоговый сигнал в диапазоне до f_m , то отсчеты должны следовать с периодом, по крайней мере, в два раза меньшим, чем период частоты f_m . Иными словами, частоту дискретизации следует выбирать так, чтобы она была, по меньшей мере, в два раза выше максимальной частоты звукового диапазона.

Чтобы из ступенчатого сигнала получить гладкую кривую, его необходимо пропустить через ФНЧ с характеристикой, аналогичной той, которую имел ФНЧ на входе АЦП (рис. 1).

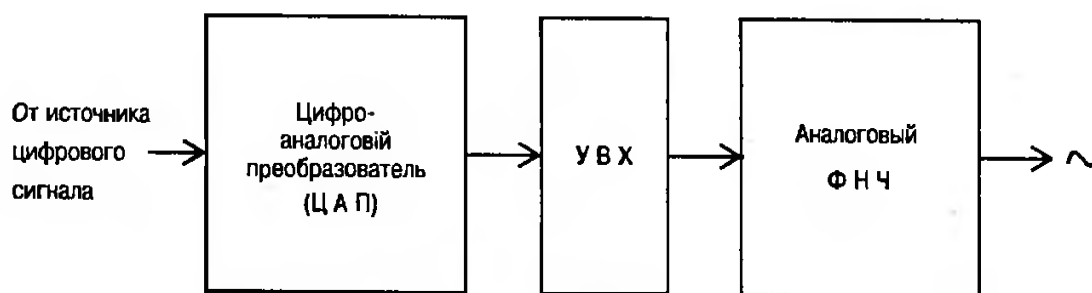


Рис. 1 Структура цифро-аналогового преобразования

Здесь также важно, чтобы спектр звукового диапазона не перекрывался спектром ближайшей боковой полосы, ибо это также приведет к появлению искажений. Для этого

характеристика ФНЧ должна иметь достаточно крутой срез — такой же, как в случае антиэлайсинг фильтра при аналого-цифровом преобразовании. Порядок такого ФНЧ должен быть не ниже 12-го. Однако построение фильтров высокого порядка связано с известными трудностями. Здесь требуется применение прецизионных пассивных элементов и высококачественных операционных усилителей с хорошей температурной и временной стабильностью. Причем, поскольку все это предназначено для использования в бытовом аппарате небольших размеров, все компоненты должны быть к тому же малогабаритными.

Кроме того, всякий фильтр высокого порядка обладает существенно нелинейной фазовой характеристикой. А это приводит к заметным на слух искажениям звуков с крутыми перепадами уровня — барабанов, тарелок, рояля и пр.

Передискретизация

Чтобы облегчить требования к фильтрации преобразованного сигнала, перед ЦАП можно разместить цифровой фильтр. Выполнить такой фильтр с нужными характеристиками значительно проще, чем аналоговый. Он может иметь достаточно высокий порядок и при этом обладать линейной фазовой характеристикой. Однако, характеристика цифрового фильтра, как и спектр цифрового сигнала, тоже имеет периодическую структуру и тоже повторяется на частотах, кратных частоте дискретизации. Поэтому, если цифровой фильтр будет работать на частоте дискретизации F_d , то подавить высокочастотные компоненты все равно не удастся.

Проблема может быть решена путем искусственного увеличения частоты дискретизации F_d в несколько раз. При этом недостающие значения сигнала вычисляются по известным значениям методами интерполяции (рис. 2).

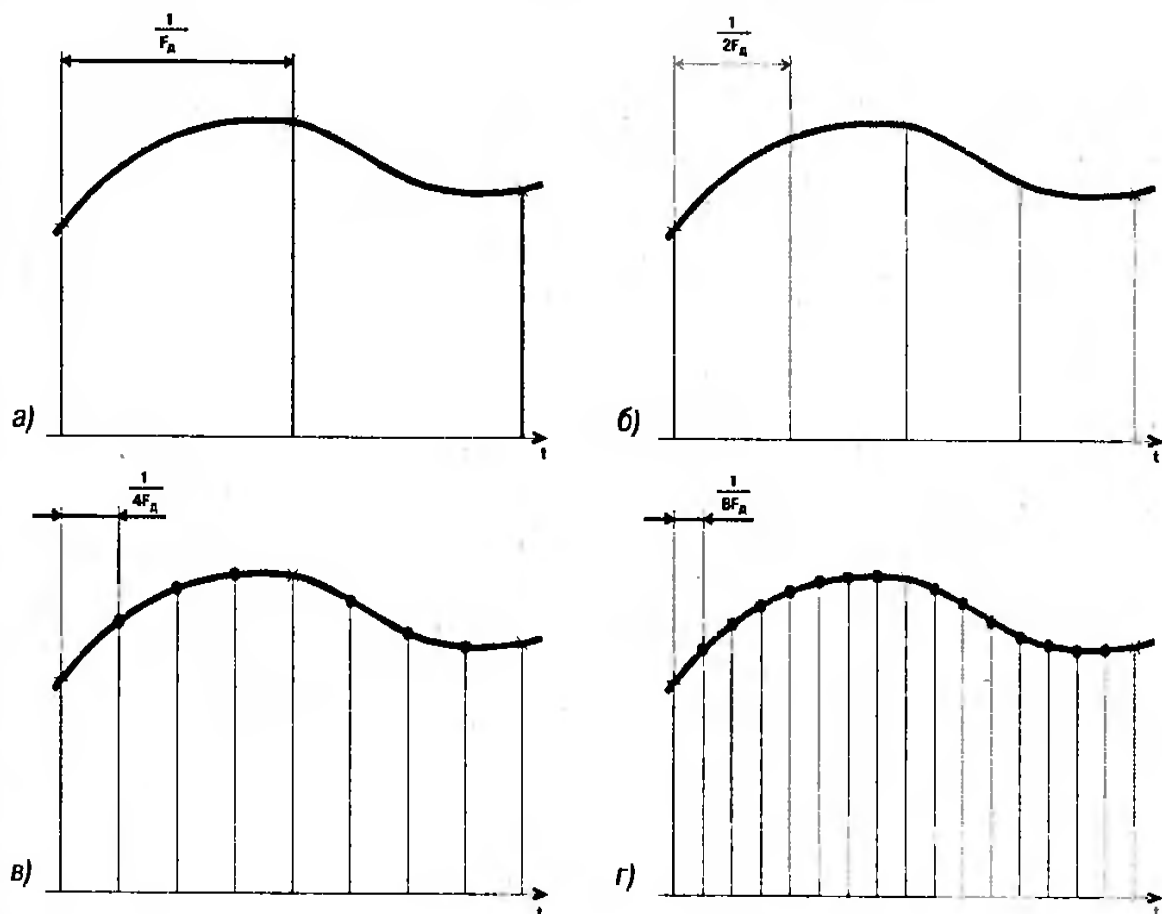


Рис. 2 Повышение частоты дискретизации с помощью интерполяции.

Передискретизация позволяет значительно снизить требования к характеристике аналогового ФНЧ. Даже удвоение F_d дает возможность сделать срез его АЧХ довольно пологим (рис. 3, а). А при увеличении частоты дискретизации в четыре, восемь и более раз, требования к аналоговому ФНЧ снижаются до вполне заурядных (рис. 3, в). За счет этого отношение сигнал/шум, а следовательно, и динамический диапазон, можно сделать даже большим, чем определяемая 16-разрядным квантованием величина в 98 дБ

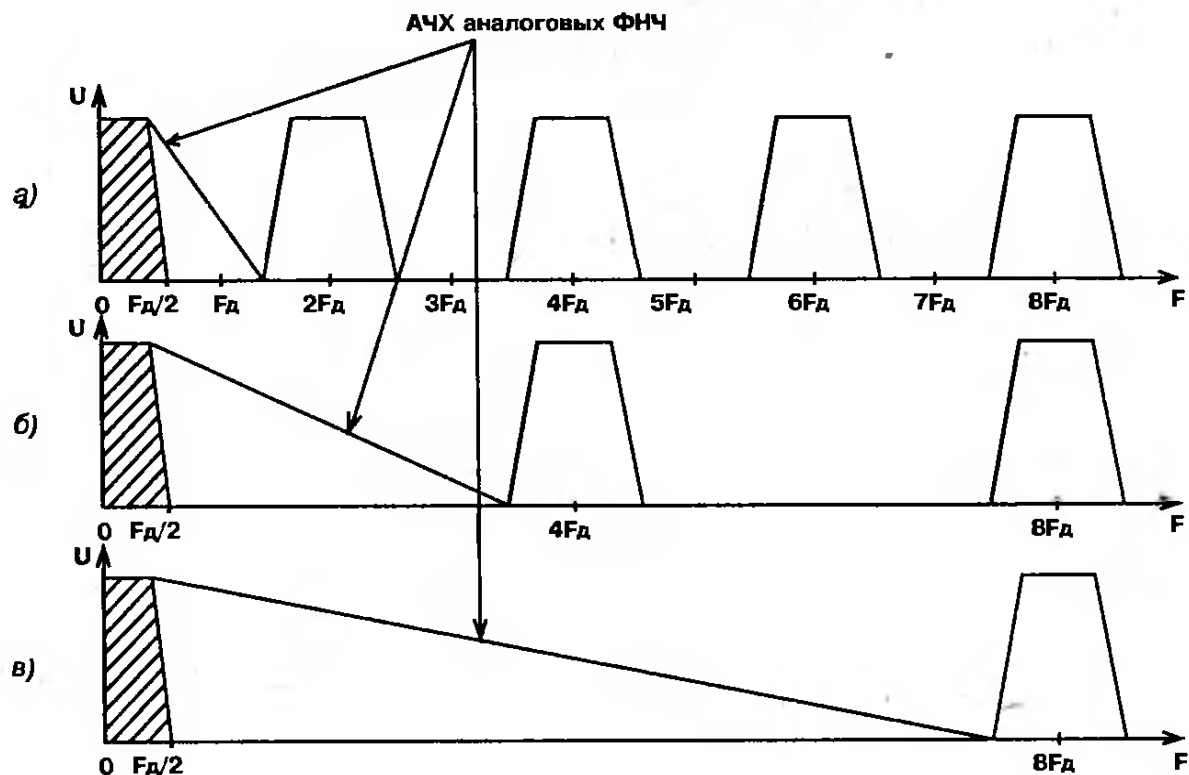


Рис. 3 Структура цифро-аналогового преобразования с интерполяцией

Схема цифро-аналогового преобразования в случае интерполяции приобретает вид, показанный на рис. 3.

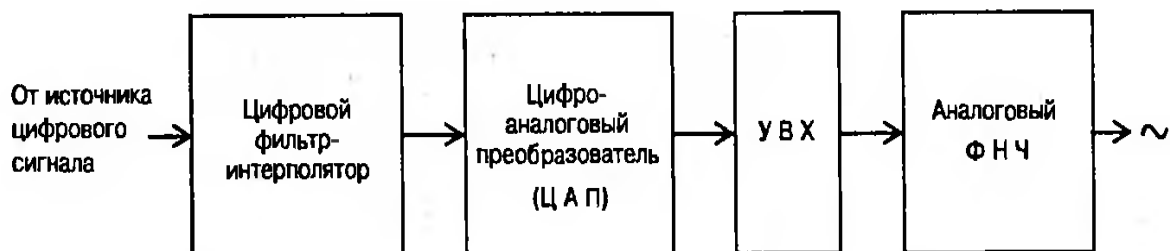


Рис. 4 Структура цифро-аналогового преобразования с интерполяцией

ЦАП для цифровой обработки звука

Широтно-импульсная модуляция. Широтно-импульсная модуляция (ШИМ) заключается в изменении ширины импульса при постоянстве частоты следования импульса. Амплитуда импульсов при этом неизменна. Регулирование значения напряжения на выходе однобитного ШИМ ЦАП осуществляется путем изменения коэффициента заполнения. Коэффициент заполнения D равен отношению длительности импульса τ к периоду импульса T :

$$D = \tau/T.$$

Среднее значение напряжения равно произведению коэффициента заполнения и амплитуды напряжения. Так, при коэффициенте заполнения $D=0,3$ и амплитуде напряжения 12В среднее значение напряжения составит $0,3 \times 12 = 3,6$ (В). При изменении коэффициента заполнения в теоретически возможных пределах от 0% до 100% напряжение будет изменяться от 0 до 12 В, т.е. широтно-импульсная модуляция позволяет регулировать напряжение в пределах от 0 до амплитуды сигнала.

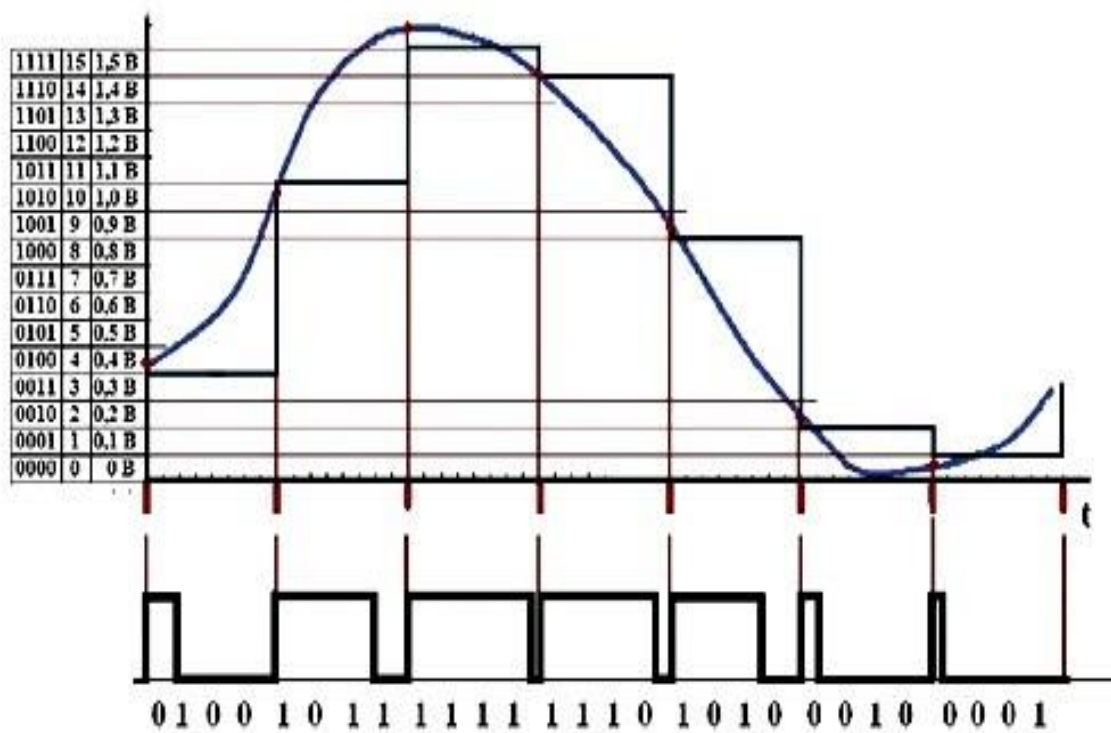


Рис. 5 ШИМ цифро-аналоговое преобразование

Сигнал с выхода ШИМ усредняется (сглаживается) посредством фильтра нижних

частот. Сглаживание сигналов ШИМ может происходить естественным путем в том случае, когда частота ШИМ превосходит время реакции регулируемого устройства.

Широтно-импульсное регулирование находит применение в схемах управления электродвигателями постоянного тока, в импульсных преобразователях, для регулирования яркости светодиодных светильников, экранов ЖК-мониторов, дисплеев в смартфонах и планшетах и т.п.

ШИМ ЦАП может быть использован для получения низкочастотных звуковых сигналов.

ЦАП с резистивной матрицей. Схема простейшего ЦАП с резистивной матрицей показана на рис. 6.

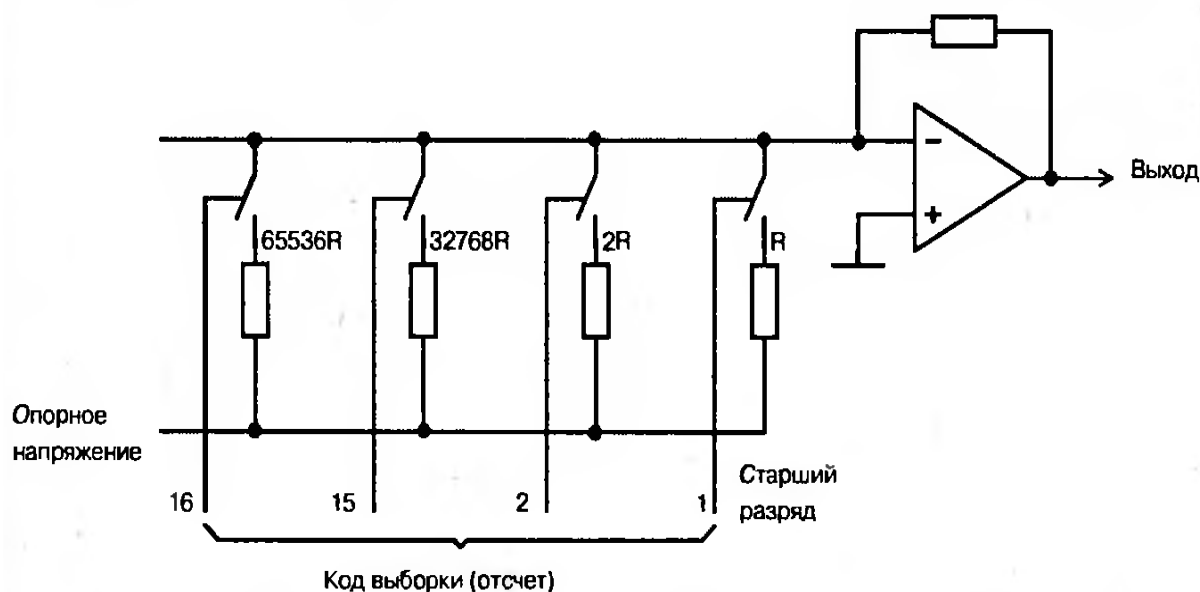


Рис. 6. Схема суммирующего ЦАП с резистивной матрицей

Он представляет собой суммирующий операционный усилитель с резистивной матрицей на входе. Число резисторов в матрице равно числу разрядов преобразуемого отсчета (на схеме — 16), каждый из которых управляет своим ключом. Если в соответствующем разряде «нуль», то ключ разомкнут, если «единица» — то замкнут. Величина сопротивления каждого последующего резистора, начиная с резистора старшего разряда, удваивается. Следовательно, ток, протекающий через эти резисторы, будет вдвое уменьшаться с уменьшением веса разряда. Выходное напряжение операционного усилителя будет пропорционально общему току и, следовательно, значению двоичного кода отсчета. Данная схема очень проста. Однако, для ЦАП высокой разрядности величина сопротивления весового резистора младшего разряда становится очень большой. Для 16-разрядного ЦАП — $65536R$. Само по себе это не страшно, но точность при этом необходима чрезвычайно высокая — $1/65536$ для 16-разрядного ЦАП. Реализовать такую точность в интегральном исполнении практически невозможно. Поэтому данная схема используется лишь для создания ЦАП низкой разрядности (до 10).

На рис. 7 показана еще одна схема ЦАП — на основе цепочки резисторов R - $2R$. Здесь используются всего два значения сопротивлений R и $2R$. Один из резисторов ($2R$)

включен последовательно с разрядным ключом, другой (R) — в суммирующую линию.

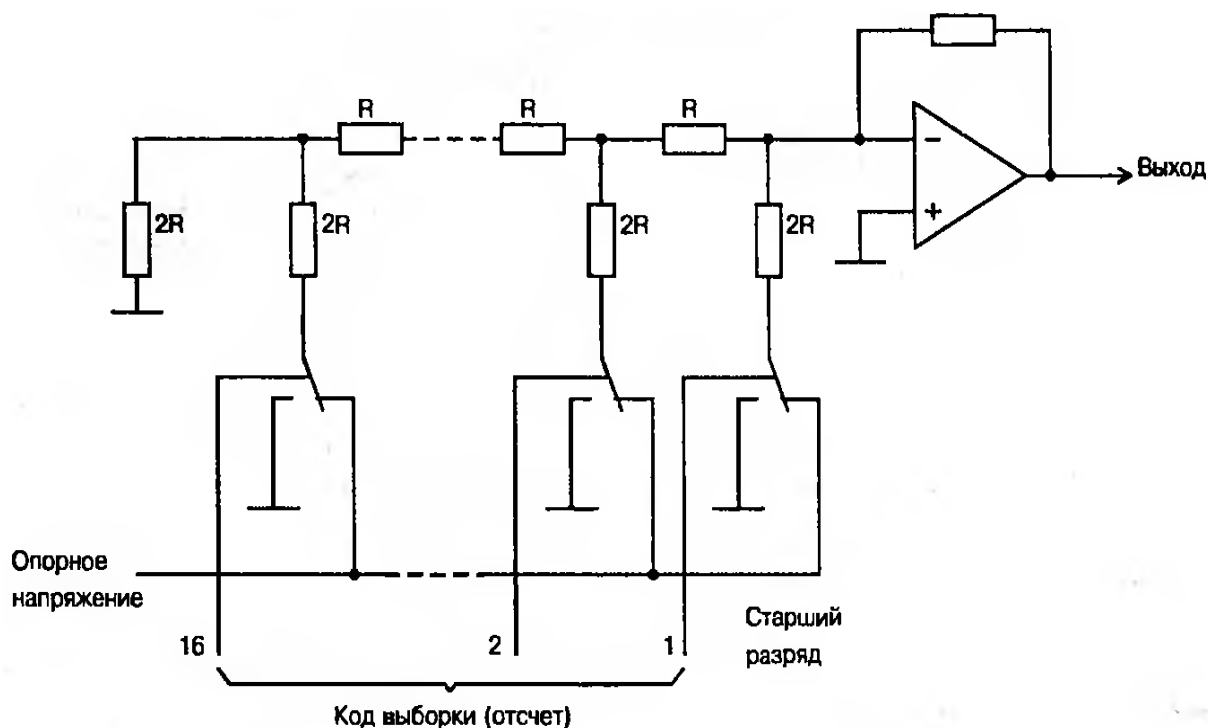


Рис. 7. Схема ЦАП на основе цепочки резисторов R - $2R$

Работает такая схема следующим образом. Если значение какого-то разряда отсчёта равно «нулю», то соответствующий ключ замыкается на «землю», если «единице» — то на шину опорного напряжения. Совокупность положений ключей обеспечивает определенное значение тока на входе операционного усилителя, который формирует на своем выходе напряжение, величина которого соответствует значению двоичного отсчета.

Эта схема проще для интегрального исполнения чем предыдущая, но и у нее есть свои недостатки. В частности, для получения точности, необходимой при реализации высокоразрядных ЦАП, коммутирующие ключи в замкнутом состоянии должны обладать сопротивлением, близким к нулевому, и бесконечным — в разомкнутом. Реально такое невозможно, и особенно — в интегральном исполнении.

Сигма-дельта ЦАП. В последнее время все более широкое распространение получают методы А/Ц и Ц/А преобразования на основе использования сигма-дельта модуляции, когда квантование осуществляется всего одним разрядом, но с частотой в десятки и сотни раз превышающей частоту Найквиста (F_H). В процессе такого преобразования анализируется не амплитуда аналогового сигнала, а направление ее изменения. Если амплитуда возрастает, то результатом преобразования будет 1, а если уменьшается — то 0. Нулевой уровень кодируется чередующимися нулями и единицами.

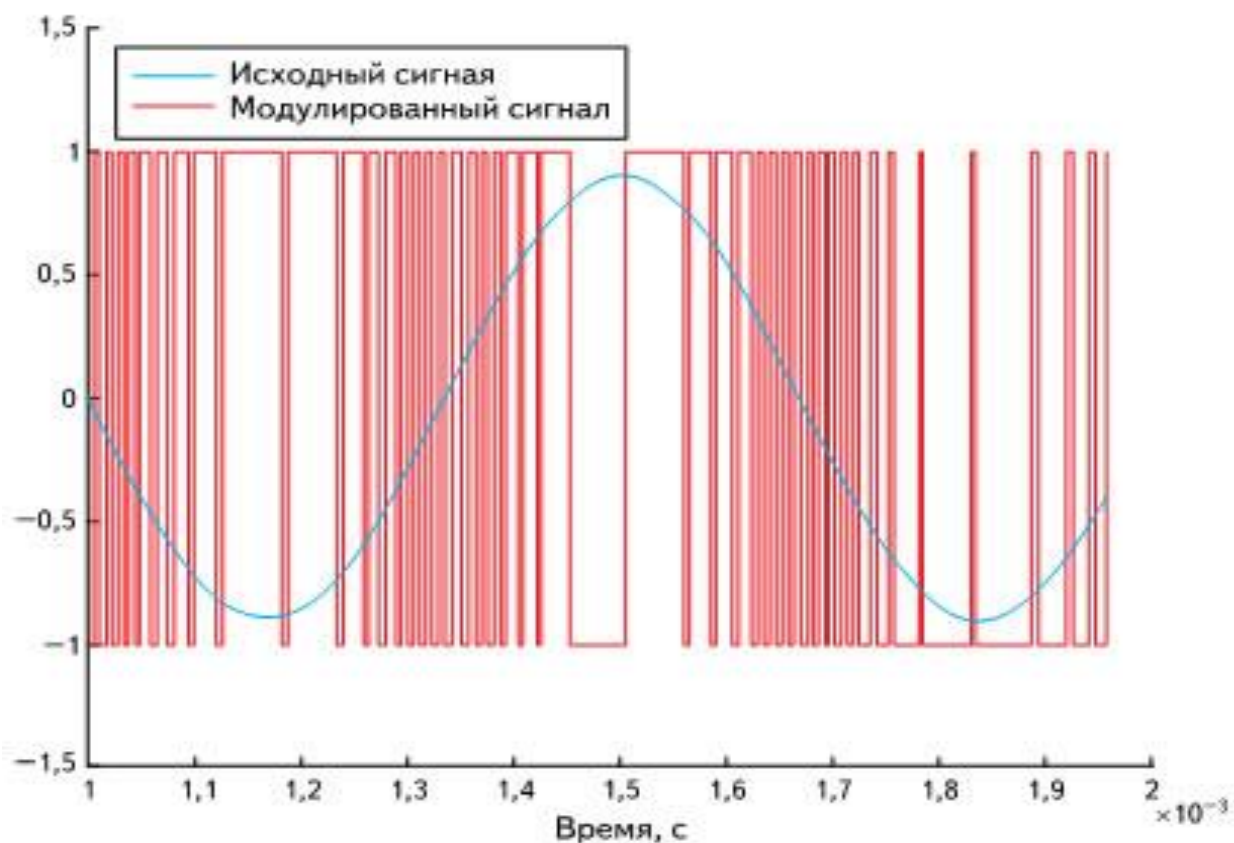


Схема блока цифро-аналогового преобразования с использованием сигма-дельта модуляции приведена на рис. 8.

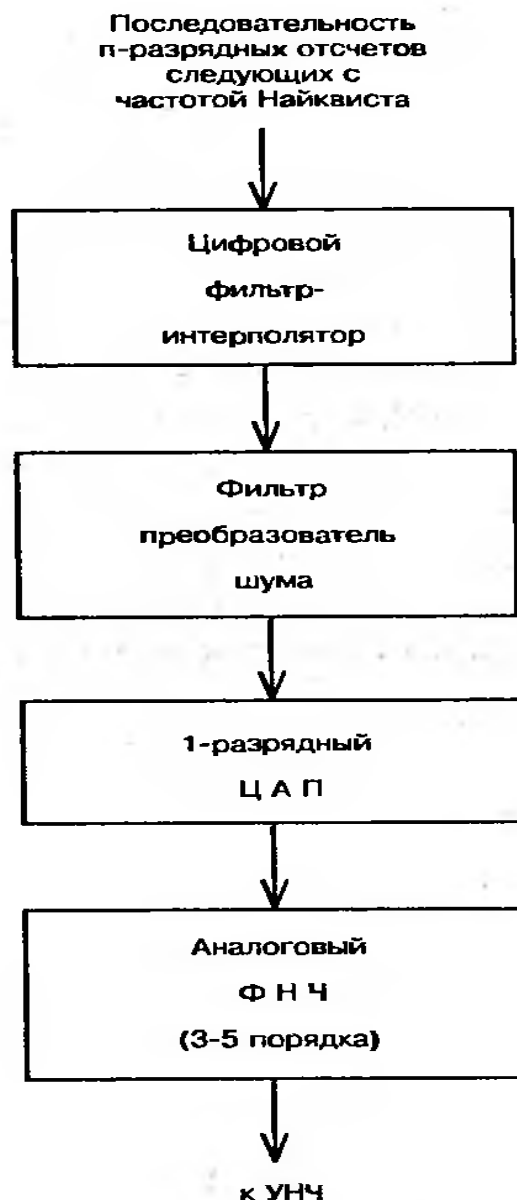


Рис. 8. Схема ЦАП с использованием сигма-дельта модулятора

Последовательность n -разрядных отсчетов, следующих с частотой Найквиста, поступает на трансверсальный фильтр-интерполятор. Здесь скорость следования отсчетов увеличивается в N раз и вычисляются промежуточные значения сигнала. Разрядность их при этом может увеличиваться, уменьшаться или оставаться прежней. После этого последовательность отсчетов поступает на рекурсивный фильтр-преобразователь шума, задачи которого те же, что и при аналого-цифровом преобразованиях — переместить часть спектра шума из основной полосы в область высших частот (рис. 9).

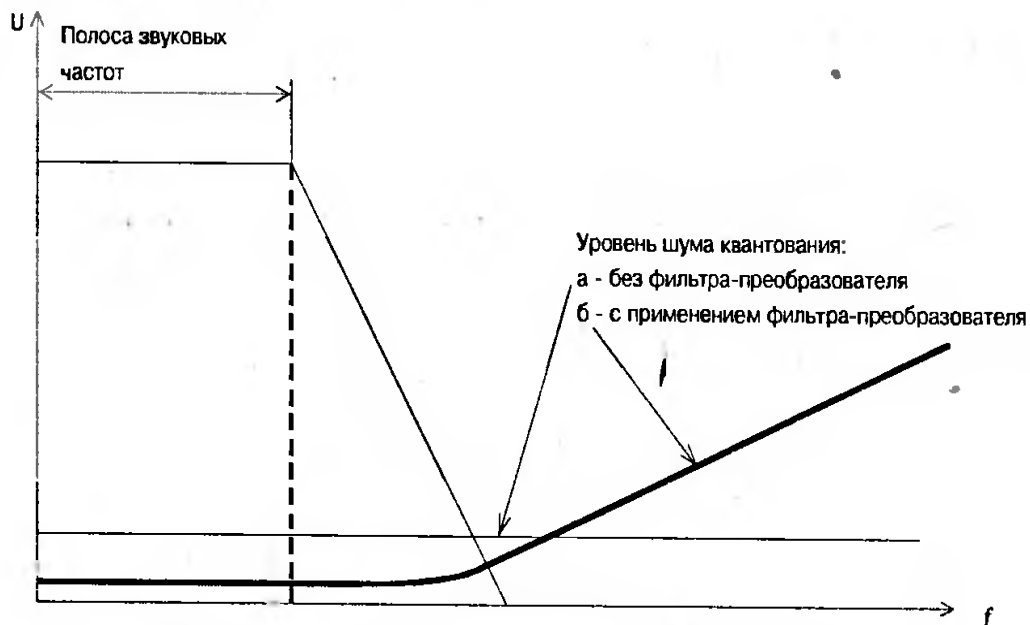


Рис. 9. Эффект применения фильтра-преобразователя

При этом разрядность отсчетов уменьшается до одного. Полученный двоичный поток подается на простой одноразрядный ЦАП, построенный на основе переключаемых емкостей, который формирует из него аналоговый сигнал. Окончательная фильтрация осуществляется аналоговым ФНЧ 3-5-го порядков.

Основным достоинством одноразрядного преобразования является простота исполнения квантователя и ЦАП, не требующих высокоточных взвешивающих элементов, которые очень сложно реализовать в интегральном исполнении.

Принцип сигма-дельта модуляции положен в основу разработанного фирмой PHILIPS метода цифро-аналогового преобразования «Bit Stream», который до сих пор широко используется ею в своих проигрывателях компакт-дисков. Алгоритм такого преобразования, реализованный в микросхеме SAA7320 и ее многочисленных модификациях, показан на рис. 10

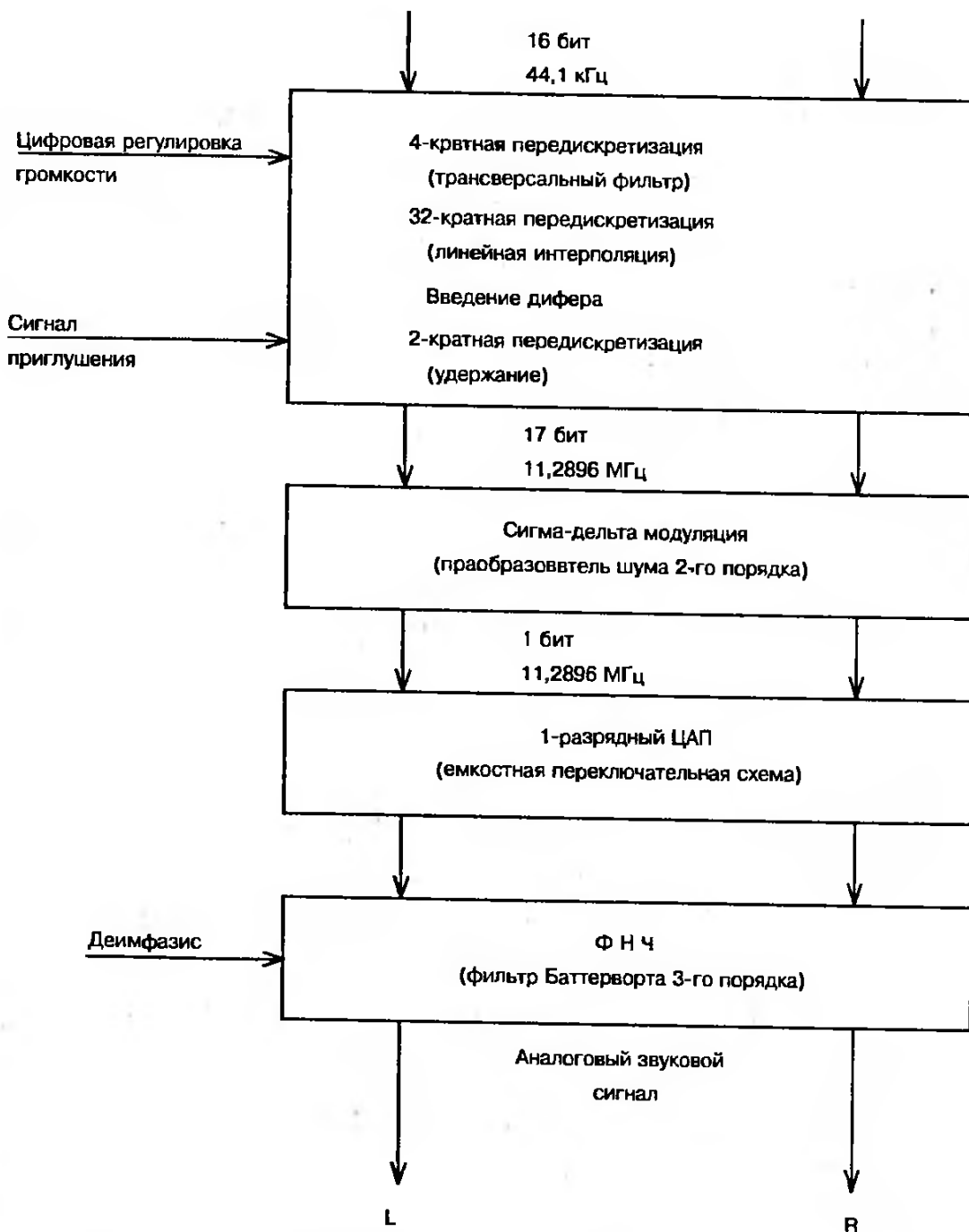


Рис. 10. Алгоритм работы системы *Bit Stream*

На первом этапе производится увеличение частоты дискретизации исходного сигнала в 256 раз ($4 \times 32 \times 2$), которая после этого становится равной 11,2396 МГц. Промежуточные значения отсчетов вычисляются вначале с помощью трансверсального фильтра с линейной фазовой характеристикой, затем с помощью линейной интерполяции и, наконец, путем удержания предыдущего значения. Здесь же к полезному сигналу подмешивается маскирующий шумоподобный сигнал (дифер), из-за чего разрядность отсчетов повышается до 17. На этом этапе, кроме передискретизации, осуществляется

еще и цифровая регулировка громкости выходного сигнала проигрывателя, а также реализуется функция приглушения, которая включается при наличии длинных последовательностей искаженных отсчетов.

После передискретизации производится преобразование спектра шума квантования с помощью сигма-дельта модулятора второго порядка. При этом число разрядов в потоке данных уменьшается до одного. В результате такой операции большая часть шумов квантования из слышимой области перемещается далеко за ее пределы.

И наконец, на последнем этапе из двоичного потока формируется аналоговый сигнал. Такое формирование осуществляется с помощью простого одноразрядного преобразователя, представляющего собой устройство с переключаемыми емкостями. Выходной каскад ЦАП является к тому же первым звеном аналогового ФНЧ.

Окончательное устранение шумов квантования осуществляется аналоговым фильтром Баттерворта третьего порядка. Здесь же производится коррекция предискажений (деимфазис), если в служебных данных присутствует сигнал о их наличии.

Лекция 5. Спектр сигнала.

Дискретное преобразование Фурье.

Как известно, звуковой сигнал в компьютере может представляться в виде некоторого набора отсчётов его амплитуд, производимых через определённые промежутки времени (период дискретизации) и представляемых некоторым количеством двоичных разрядов (разрядность выборки). Такое представление (во временной области) удобно для хранения звукового сигнала и его преобразования обратно в непрерывный сигнал. Однако, в некоторых случаях при выполнении обработки звукового сигнала (например, фильтрация) целесообразно перейти к спектру сигнала.

Частотный спектр сигнала — это распределение энергии сигнала по частотам. Спектр бывает амплитудный и фазовый.

Доказано, что если некоторая периодическая функция с периодом $2T$ на интервале $[-T, T]$ удовлетворяет условиям Дирихле (непрерывна и имеет конечное число экстремумов и точек разрыва I рода), то она может быть представлена в виде суммы ряда Фурье (разложена в ряд Фурье):

$$f(x) = a_0 + \sum_{n=1}^{\infty} \left(a_n \cos \frac{n\pi x}{T} + b_n \sin \frac{n\pi x}{T} \right)$$

Для определения коэффициентов ряда Фурье справедливы следующие формулы:

$$a_n = \frac{1}{T} \int_{-T}^T f(x) \cos \frac{n\pi x}{T} dx$$

$$b_n = \frac{1}{T} \int_{-T}^T f(x) \sin \frac{n\pi x}{T} dx$$

Если раскладываемая функция является чётной ($f(-x) = f(x)$), то ряд Фурье состоит только из косинусов, т. е. все коэффициенты при синусах равны 0. Если раскладываемая

функция является нечётной ($f(-x) = -f(x)$), то ряд Фурье состоит только из синусов, т. е. все коэффициенты при косинусах равны 0. В общем случае, коэффициенты при синусах и косинусах не равны 0.

Таким образом, любую периодическую функцию, удовлетворяющую условиям Дирихле, можно разложить в ряд Фурье, тем самым представляя её в виде суммы синусов и косинусов.

Непериодические (апериодические) сигналы, например, одиночный импульс, разложить в *конечный* ряд Фурье невозможно. Но если апериодический сигнал ограничить во времени (например, интервалом $[-T, T]$), то его можно дополнить такими же сигналами, но сдвинутыми относительно друг друга на $2T$, и получить периодический сигнал.

Спектр дискретного периодического сигнала может быть рассчитан при помощи дискретного преобразования Фурье (ДПФ).

Дискретное преобразование Фурье имеет вид:

$$X_k = \frac{1}{N} \sum_{i=0}^{N-1} x_i e^{-j\frac{2\pi ki}{N}} = \frac{1}{N} \sum_{i=0}^{N-1} x_i \left[\cos \frac{2\pi ki}{N} - j \sin \frac{2\pi ki}{N} \right], \quad k = 0, 1, \dots, N-1$$

Обратите внимание, что косинусоидальные и синусоидальные компоненты в уравнении могут быть выражены в полярных или прямоугольных координатах, связь между которыми определяется формулой Эйлера:

$$e^{-j\varphi} = \cos\varphi - j\sin\varphi$$

После этого преобразования звуковой сигнал будет представлен в виде двух массивов:
- массив, содержащий значения косинусоидальных составляющих,

$$ReX_k = \frac{1}{N} \sum_{i=0}^{N-1} x_i \cos \frac{2\pi ki}{N}, \quad k = 0, 1, \dots, N-1$$

- массив, содержащий значения синусоидальных составляющих.

$$ImX_k = -\frac{1}{N} \sum_{i=0}^{N-1} x_i \sin \frac{2\pi ki}{N}, \quad k = 0, 1, \dots, N-1$$

Такие обозначения введены в силу комплексного представления преобразования Фурье. При этом действительной части соответствуют косинусы, а мнимой - синусы. Это не должно вводить Вас в заблуждение - элементы этих массивов являются действительными числами и коэффициенты при синусах и косинусах являются действительными числами. Массивы ReX и ImX составляют так называемый частотный домен (frequency domain), в то время как исходная выборка называется временным доменом (time domain).

Амплитудный спектр звукового сигнала вычисляется по формуле:

$$Abs[X_k] = \sqrt{ReX_k^2 + ImX_k^2}, \quad k = 0, 1, \dots, N-1$$

Фазовый спектр звукового сигнала вычисляется по формуле:

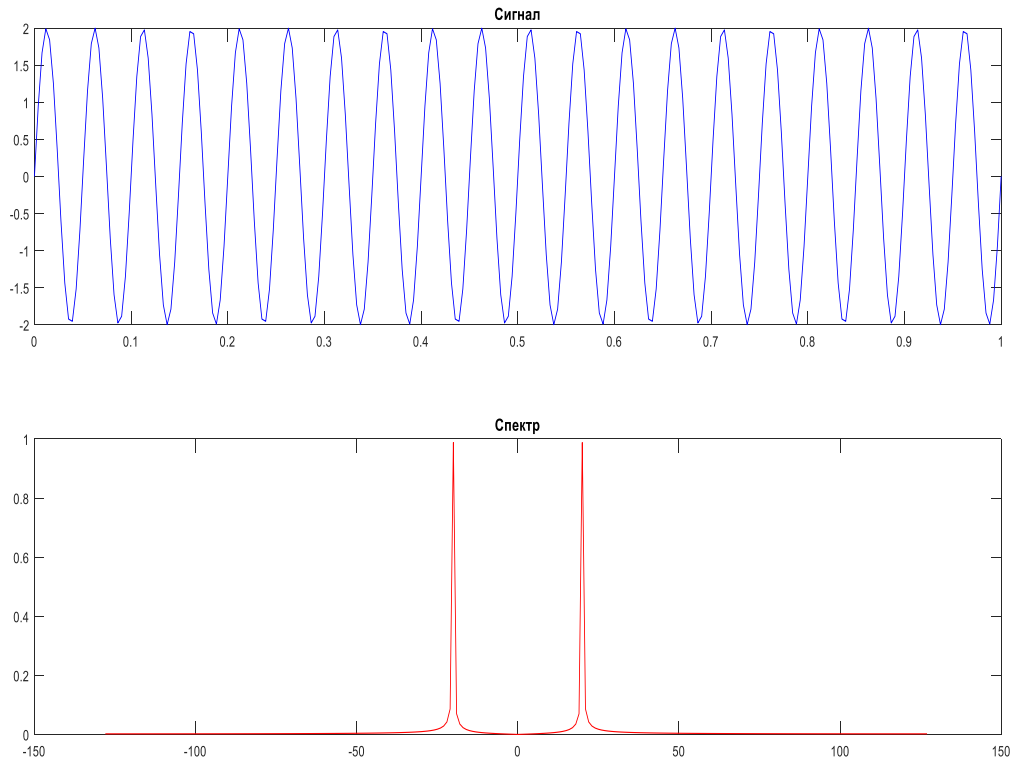
$$\varphi[X_k] = \tan^{-1} \frac{ImX_k}{ReX_k}, \quad k = 0, 1, \dots, N-1$$

Фактически, дискретное преобразование Фурье позволяет представить дискретную периодическую функцию в виде конечного числа частот с определёнными значениями амплитуды и фазы (раскладывает функцию в её спектр). Для определения амплитуд и фаз частотных составляющих сигнала, в дискретном преобразовании Фурье используются базисные функции синуса и косинуса. Спектр частот в дискретном преобразовании Фурье определяется из амплитуд синусов и косинусов, с частотами повторения в исследуемой выборке от 0 до $N/2$ раз, где N - количество элементов выборки.

Преобразуемый сигнал может быть, как действительным (мнимая часть равна нулю), так и комплексным. Если сигнал действительный, то спектры для отрицательных и положительных частот совпадают и для анализа достаточно одного из них (см. пример 1). Если сигнал комплексный, то спектры для отрицательных и положительных частот в общем случае различны (знак частоты имеет значение – см. примеры 2 и 3).

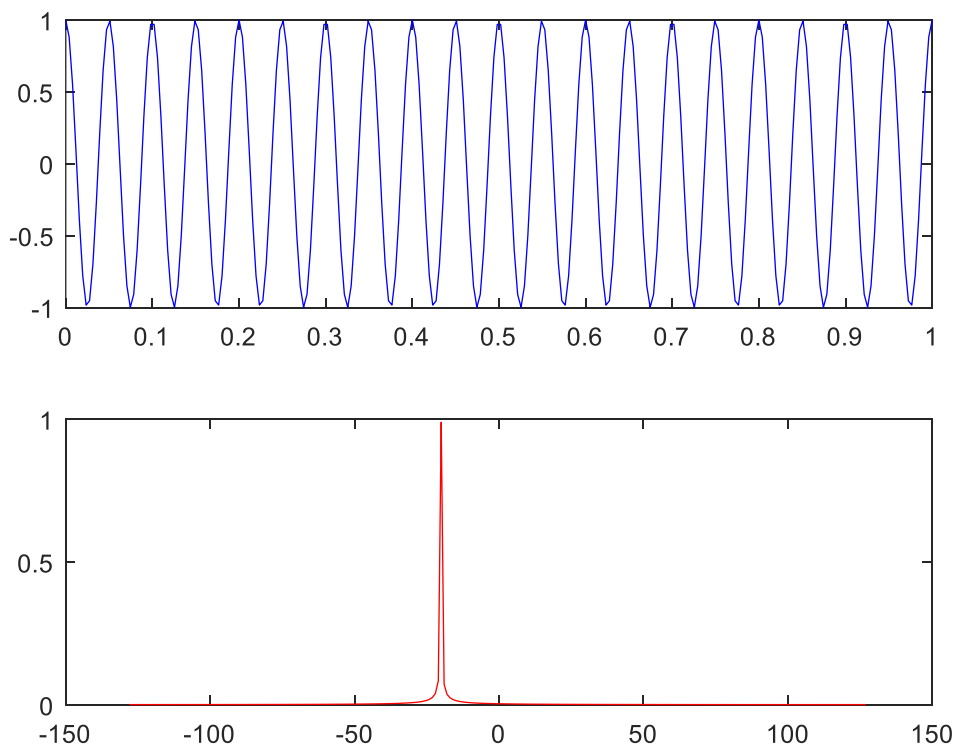
Примеры 1: Амплитудный спектр действительного сигнала частотой 20 Гц

```
N=256;%Число точек
T=1;%Интервал времени,с
dt=T/(N);
Nyq=N/(2*T)%Частота Найквиста, Гц
df=1/T;%Шаг частоты
f=20;%Частота сигнала
w=2*pi*f;
t=linspace(0,T,N);%Задание вектора времени
s=2*sin(w*t./T);%Вычисление дискретной функции
F=fft(s)/N;%Вычисление преобразования Фурье
subplot(211);
plot(t,real(s),'b-');%Отрисовка исходной функции
F1=F(1:N/2+1);%Выделение первых N/2+1 (положительные частоты)
F2=F(N/2+1:N);%Выделение спектра отрицательных частот
F=[F2,F1];%Объединение спектра
%Вычисление вектора частот
nu=-Nyq+df*(0:N);
subplot(212);
plot(nu(1:N),abs(F(1:N)),'r-')
```



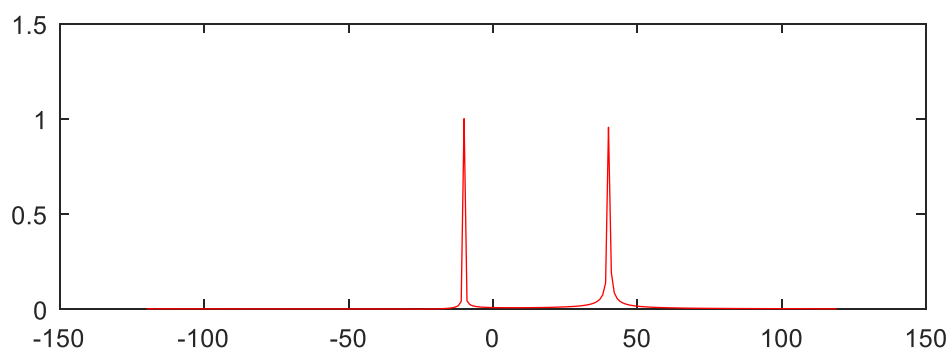
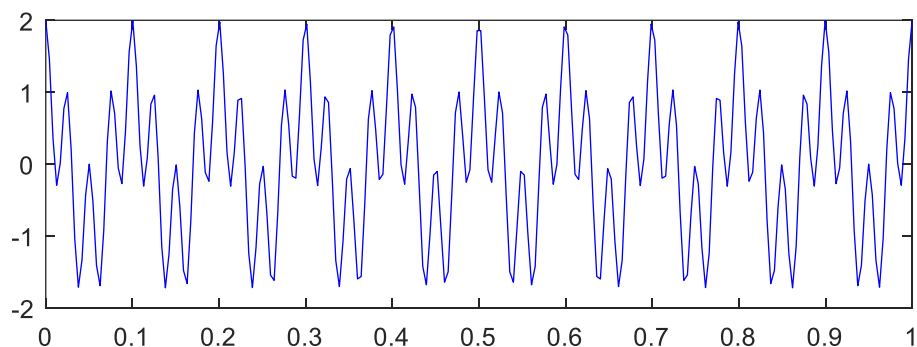
Примеры 2: Спектр комплексного сигнала частотой -20 Гц

```
N=256;%Число точек
T=1;%Интервал времени,с
dt=T/(N);
Nyq=N/(2*T)%Частота Найквиста, Гц
df=1/T;%Шаг частоты
f=-20;%Частота сигнала
w=2*pi*f;
t=linspace(0,T,N);%Задание вектора времени
s=2*(cos(w*t./T)+i*sin(w*t./T));%Вычисление дискретной функции
F=fft(s)/N;%Вычисление преобразования Фурье
subplot(211);
plot(t,real(s),'b-');%Отрисовка исходной функции
F1=F(1:N/2+1);%Выделение первых N/2+1 (положительные частоты)
F2=F(N/2+1:N);%Выделение спектра отрицательных частот
F=[F2,F1];%Объединение спектра
%Вычисление вектора частот
nu=-Nyq+df*(0:N);
subplot(212);
plot(nu(1:N),abs(F(1:N)),'r-')
```



Примеры 3: Спектр двух тонального комплексного сигнала с частотами -10 Гц и 40 Гц

```
N=240;%Число точек
T=1;%Интервал времени,с
dt=T/(N);
Nyq=N/(2*T)%Частота Найквиста, Гц
df=1/T;%Шаг частоты
f1=-10;%Частота сигнала 1
w1=2*pi*f1;
f2=40;%Частота сигнала 1
w2=2*pi*f2;
t=linspace(0,T,N);%Задание вектора времени
s=1*(cos(w1*t./T)+i*sin(w1*t./T))+1*(cos(w2*t./T)+i*sin(w2*t./T));%Вычисление дискретной функции
F=fft(s)/N;%Вычисление преобразования Фурье
subplot(211);
plot(t,real(s),'b-');%Отрисовка исходной функции
F1=F(1:N/2+1);%Выделение первых N/2+1 (положительные частоты)
F2=F(N/2+1:N);%Выделение спектра отрицательных частот
F=[F2,F1];%Объединение спектра
%Вычисление вектора частот
nu=-Nyq+df*(0:N);
subplot(212);
plot(nu(1:N),abs(F(1:N)),'r-')
```



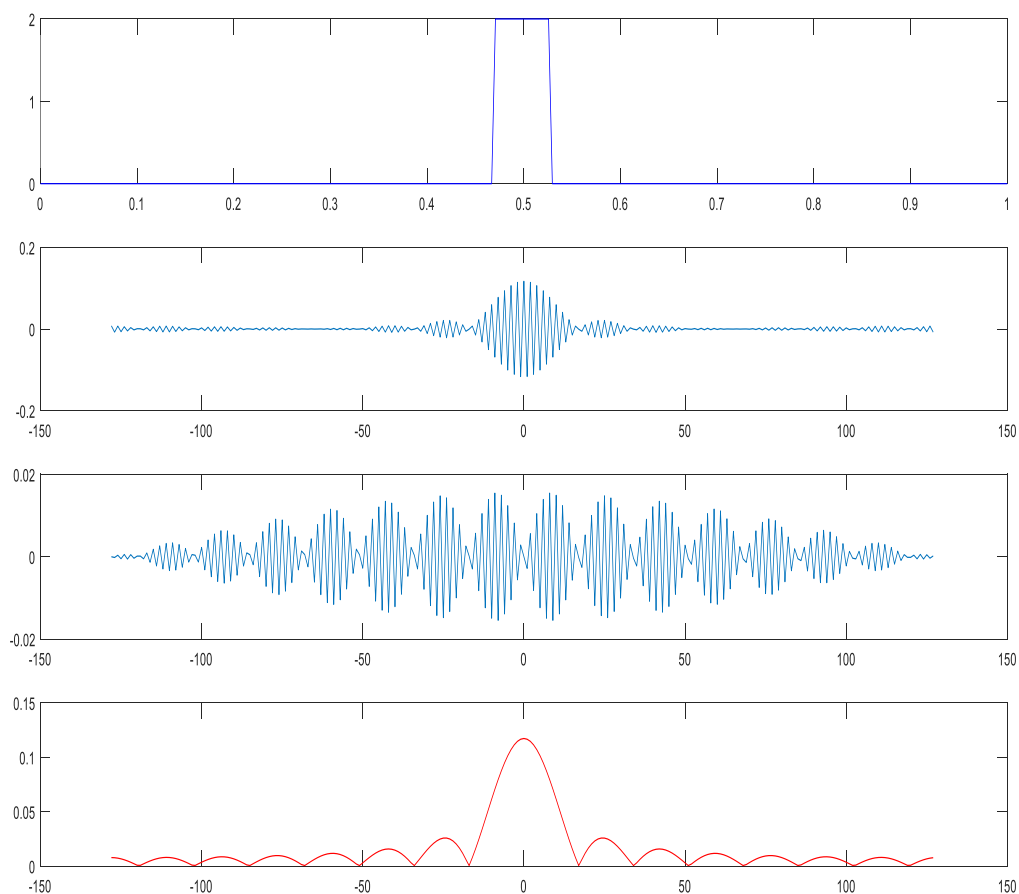
Пример 5. Спектр прямоугольного импульса

```
%%Прямоугольный импульс
N=256;%Число точек
T=1;%Интервал времени,с
dt=T/(N);
Nyq=N/(2*T)%Частота Найквиста, Гц
df=1/T;%Шаг частоты
f=50;%Частота сигнала
w=-2*pi*f;
t=linspace(0,T,N);%Задание вектора времени

for x=1:1:N;

s(1,x)=0;
if x > N/2 - 8 & x < N/2 + 8;%Вычисление дискретной функции
    s(1,x)=2;
end

end;
F=fft(s)/N;%Вычисление преобразования Фурье
subplot(411);
plot(t,s,'b-');%Отрисовка исходной функции
F1=F(1:N/2+1);%Выделение первых N/2+1 (положительные частоты)
F2=F(N/2+1:N);%Выделение спектра отрицательных частот
F=[F2,F1];%Объединение спектра
%Вычисление вектора частот
nu=-Nyq+df*(0:N);
subplot(412);
plot(nu(1:N),real(F(1:N)))
subplot(413);
plot(nu(1:N),imag(F(1:N)))
subplot(414);
plot(nu(1:N),abs(F(1:N)),'r-') Красный – амплитудный спектр
```



Примеры 6: Спектр дельта-импульса (длительность - один отчет)

```

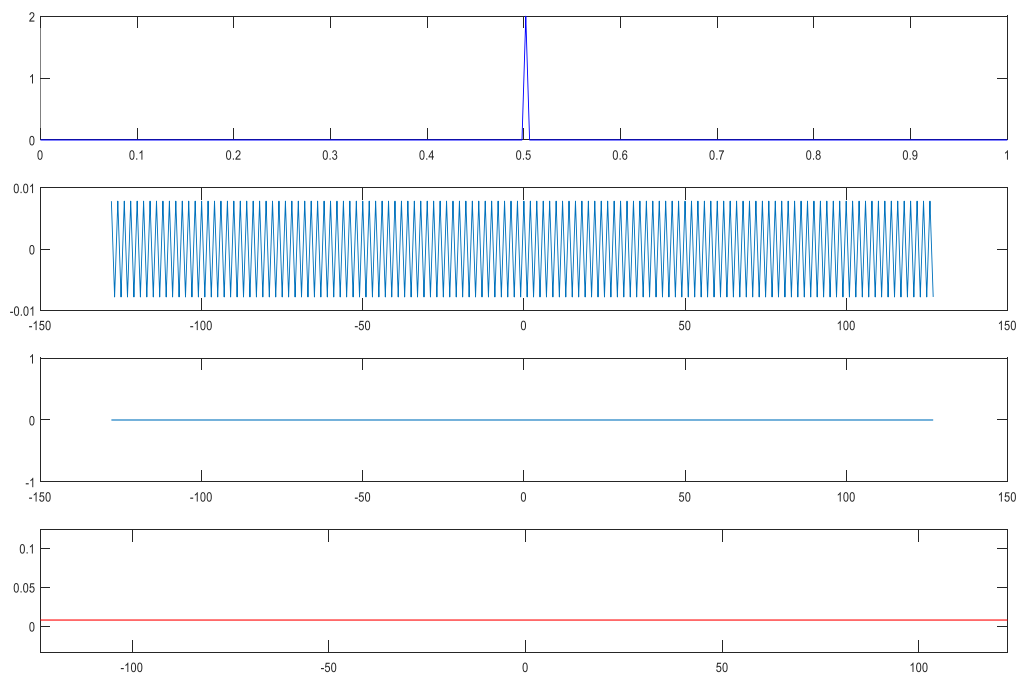
%%%Одиночный импульс
N=256;%Число точек
T=1;%Интервал времени,с
dt=T/(N);
Nyq=N/(2*T)%Частота Найквиста, Гц
df=1/T;%Шаг частоты
f=50;%Частота сигнала
w=-2*pi*f;
t=linspace(0,T,N);%Задание вектора времени

for x=1:1:N;

s(1,x)=0;
if x > N/2 & x < N/2 + 2;%Вычисление дискретной функции
    s(1,x)=2;
end

end;
F=fft(s)/N;%Вычисление преобразования Фурье
subplot(411);
plot(t,s,'b-');%Отрисовка исходной функции
F1=F(1:N/2+1);%Выделение первых N/2+1 (положительные частоты)
F2=F(N/2+1:N);%Выделение спектра отрицательных частот
F=[F2,F1];%Объединение спектра
%Вычисление вектора частот
nu=-Nyq+df*(0:N);
subplot(412);
plot(nu(1:N),real(F(1:N)))
subplot(413);
plot(nu(1:N),imag(F(1:N)))
subplot(414);
plot(nu(1:N),abs(F(1:N)),'r-') Красный - амплитудный спектр

```



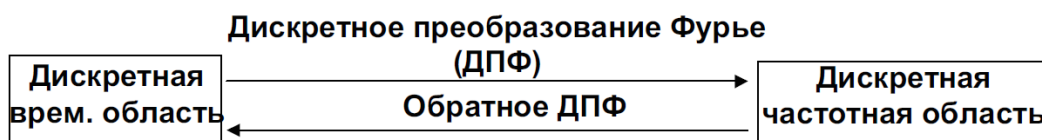
Обратное дискретное преобразование Фурье

Обратное дискретное преобразование Фурье (ОДПФ) имеет вид:

$$x_n = \sum_{i=0}^{N-1} X_i e^{j\frac{2\pi ni}{N}} = \sum_{i=0}^{N-1} X_i \left[\cos \frac{2\pi ni}{N} + j \sin \frac{2\pi ni}{N} \right], n = 0, 1, \dots, N-1$$

ОДПФ - из частотных составляющих позволяет сформировать исходный сигнал. Такой процесс преобразования называется синтезом или обратным преобразованием Фурье. Заметим, что формулы обратного преобразования аналогичны формулам прямого преобразования. Это свойство является очень важным и называется двойственностью преобразования Фурье. Свойство двойственности позволяет объяснить следующий факт: единичный импульс во временном домене (единичное значение одной выборки при нулевых значениях остальных) соответствует синусоиде в частотном домене и наоборот (см. примеры 2 и 6).

Применение дискретного преобразования Фурье (ДПФ)



- Цифровой спектральный анализ
 - ◆ Спектральные анализаторы
 - ◆ Обработка речевого сигнала
 - ◆ Обработка сигналов изображения
 - ◆ Распознавание образов
- Проектирование фильтров
 - ◆ Вычисление импульсной характеристики по частотной
 - ◆ Вычисление частотной характеристики по импульсной
- Быстрое преобразование Фурье (БПФ) – просто алгоритм для эффективного вычисления дискретного преобразования Фурье

Быстрое преобразование Фурье

ДПФ и ОДПФ для своей реализации требуют выполнения N^2 умножений комплексных чисел. ДПФ может быть сильно упрощено, если использовать свойства симметрии и периодичности коэффициентов. Результатом переработки выражений для ДПФ является быстрое преобразование Фурье (БПФ), которое требует только $\frac{N \log_2 N}{2}$ умножений комплексных чисел, т.е. количество умножений комплексных чисел уменьшается в $\frac{2N}{\log_2 N}$ раз. Вычислительная эффективность БПФ по сравнению с ДПФ становится весьма существенной, когда количество точек БПФ увеличивается до нескольких тысяч. Например, при $N=4096$ выигрыш – в 682 раза.

Очевидно, что БПФ вычисляет все компоненты выходного спектра (или все, или ни одного!). Если необходимо рассчитать только несколько точек спектра, ДПФ может оказаться более эффективным. Вычисление одного выходного отсчета спектра с использованием ДПФ требует только N умножений с комплексными числами.

Лекция 6. Фильтрация звука

Под фильтрацией применительно к обработке аудиоданных понимается процесс преобразования электрического звукового сигнала частотноизбирательными устройствами с целью изменения спектрального состава (тембра) сигнала.

Задачами такой обработки могут быть:

- амплитудно-частотная коррекция сигнала (усиление или ослабление отдельных частотных составляющих)
- подавление шумов
- улучшение качества сигнала (например, устранение или снижение помех),
- извлечение из сигналов информации или разделение нескольких сигналов.

Например, если микрофон, акустическая система или еще какой-либо элемент звукового тракта имеют неравномерную амплитудно-частотную характеристику, то с помощью фильтров эти неравномерности могут быть сглажены. Или, скажем, если в результате анализа спектра выяснилось, что в некоторой области частот сосредоточена в основном энергия помех, а энергии сигнала совсем немного, то посредством фильтрации все колебания в этом диапазоне частот можно подавить.

Фильтрация сводится к умножению спектральных коэффициентов (спектра) сигнала на соответствующие значения передаточной функции фильтра. Сигнал описывается совокупностью амплитудного и фазового спектров (АС и ФС), а фильтры - амплитудночастотными и фазочастотными характеристиками (АЧХ и ФЧХ). АЧХ представляет собой зависимость коэффициента передачи фильтра от частоты, ФЧХ отражает сдвиг фазы выходного сигнала по отношению ко входному в зависимости от частоты. В этом случае фильтрация эквивалентна перемножению АС на АЧХ и алгебраическому сложению ФС с ФЧХ.

Типы фильтров

В зависимости от вида АЧХ различают:

- Фильтры нижних частот (ФНЧ) (Low Pass)
- Фильтры верхних частот (ФВЧ) (High Pass)
- Полосовые фильтры (Band Pass)
- Заградительные (режекторные) фильтры (Band Stop)

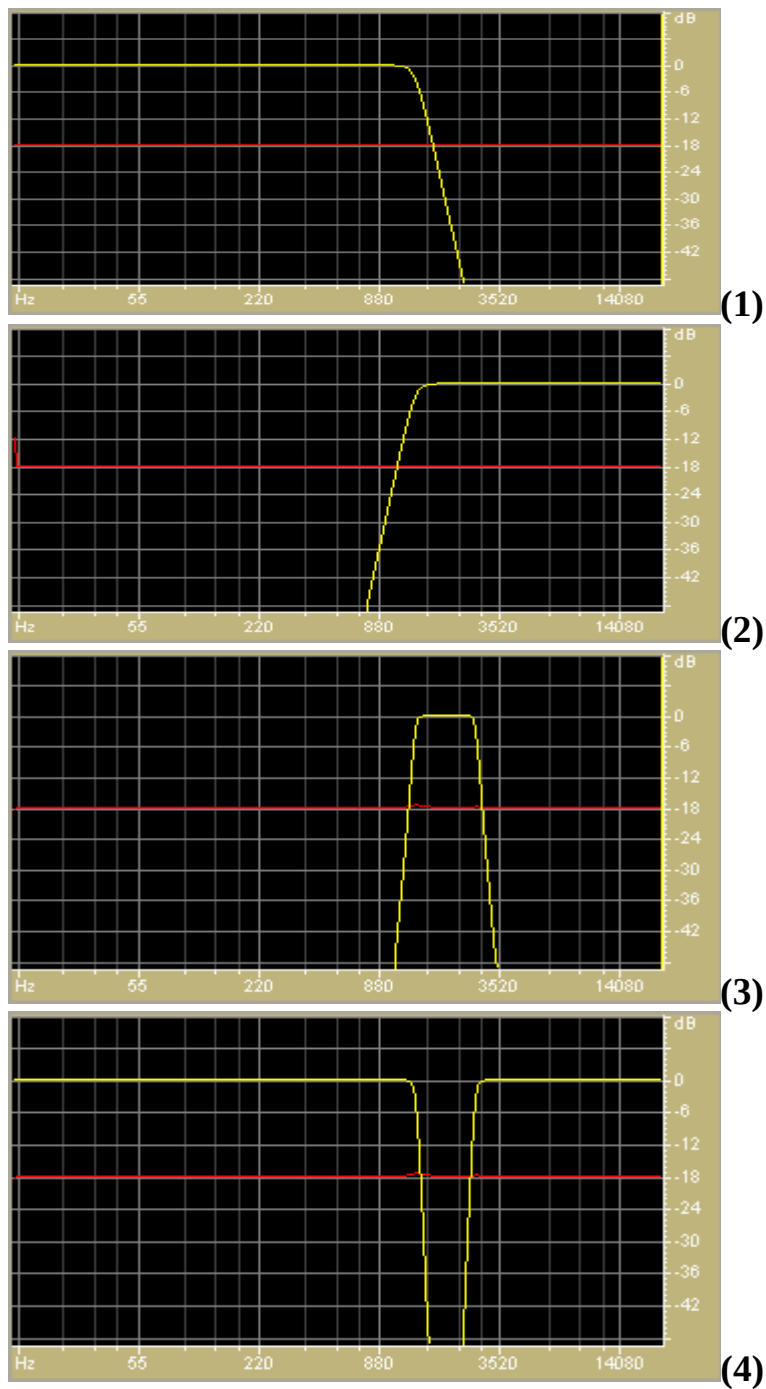


Рис. 1 - 4. АЧХ фильтров нижних частот, верхних частот, полосового и режекторного соответственно.

Тот участок АЧХ, где коэффициент передачи велик, приблизительно соответствует *полосе пропускания фильтра*. В полосе задерживания (подавления), напротив, этот коэффициент должен быть минимальным.

Реальные фильтры низких и высоких частот характеризуются такими параметрами, как *частота среза* (разграничивает полосы пропускания и подавления), неравномерность характеристик в полосе пропускания, крутизна ската характеристики в области перехода от полосы пропускания к полосе задерживания. Для полосового фильтра добавляются еще два параметра - ширина полосы пропускания (подавления) и добротность, то есть отношение центральной частоты фильтра к ширине полосы пропускания (подавления).

Фильтры в зависимости от способа реализации делятся на *аналоговые* и *цифровые*.

Аналоговые фильтры изготавливают на основе либо колебательных звеньев, состоящих из катушек индуктивности и конденсаторов, либо так называемых гираторов, представляющих собой операционные усилители, охваченные определенными обратными связями.

Цифровые фильтры

Цифровой фильтр — обозначают определенную аппаратную или программную процедуру, реализацию алгоритма фильтрации. В цифровых фильтрах используются оцифрованные аналоговые сигналы или просто хранящиеся в памяти компьютера числа, представляющие некоторые переменные. По сути, фильтр — это система или сеть, избирательно меняющая форму сигнала (амплитудно-частотную или фазово-частотную характеристику).

Упрощенная схема цифрового фильтра реального времени с аналоговым входом и выходом приведена на рисунке ниже. Аналоговый входной сигнал периодически выбирается и конвертируется в набор цифровых отсчетов, цифровой процессор производит фильтрацию, преобразует входную последовательность в выходную, согласно вычислительному алгоритму фильтра. ЦАП конвертирует отфильтрованный цифровым образом выходную отсчеты в аналоговый сигнал, который затем проходят

аналоговую фильтрацию для сглаживания и устранения нежелательных высокочастотных компонентов.



Упрощенная схема цифрового фильтра реального времени с аналоговым входом и выходом

Обобщенный алгоритм цифровой фильтрации имеет следующий вид:

$$y(n) = \sum_{i=0}^K b_i x(n-i) + \sum_{j=1}^M a_j y(n-j)$$

где $x(i)$ – отсчёты входной последовательности

$y(i)$ – отсчёты выходной последовательности

b_i, a_j – коэффициенты фильтра, определяются на основе требуемой АЧХ.

Порядок фильтра - $M+K$. Если $M=0$, то фильтр нерекурсивный и имеет конечную импульсную характеристику (КИХ). Если $M>0$, то фильтр рекурсивный и имеет бесконечную импульсную характеристику (БИХ).

Цифровая фильтрация при помощи ДПФ и ОДПФ

Цифровая фильтрация при помощи ДПФ и ОДПФ предполагает следующий порядок действий: берем преобразование Фурье от входного сигнала, умножаем результат на АЧХ фильтра (комплексный коэффициент передачи), выполняем обратное преобразование Фурье, результатом которого является выходной сигнал.

Это возможно только, если выборки сигнала (фонограмма) конечной длительности.

Реальном дискретном времени такой порядок действий выполнить невозможно.

В этом случае используется выборки необходимо обрабатывать преобразованием Фурье с перекрытием. То есть, каждая последующая выборка должно содержать часть предыдущей. В идеальном случае выборки должны перекрываться на (2^{n-1}) отсчетов, но это требует огромных вычислительных затрат. На практике, достаточно трехчетвертного ($2^{n-2^{(n-2)}}$), половинного ($2^{(n-1)}$) и даже четвертного перекрытия ($2^{(n-2)}$).

Цифровые фильтры играют важную роль в цифровой обработке сигналов. По сравнению с аналоговыми фильтрами они предпочтительны во множестве областей (например, сжатие данных, биомедицинская обработка сигналов, обработка речи, обработка изображений, передача данных, цифровое аудио, телефонное эхоподавление), так как обладают рядом преимуществ, часть из которых описана ниже.

- цифровые фильтры могут иметь характеристики, получить которые на аналоговых фильтрах практически невозможно, например, линейную фазовую характеристику;

- в отличие от аналоговых, производительность цифровых фильтров не зависит от изменений среды, например, от колебаний температуры. Таким образом, цифровые фильтры не требуют периодической калибровки;

- обрабатываемые данные представлены в цифровом виде, поэтому их просто сохранить для последующего использования;

- на практике точность, которой можно добиться при использовании аналоговых фильтров, ограничена; например, затухание в полосе подавления нельзя поднять выше 60-70 дБ (если использовать стандартные аналоговые компоненты). Точность цифровых фильтров ограничена только используемой разрядностью данных;

- цифровые фильтры могут использоваться при очень низких частотах, характерных, например, для многих биомедицинских приложений, где применять аналоговые фильтры невозможно. Кроме того, цифровые фильтры могут использоваться в большом диапазоне частот, для чего достаточно просто менять частоту дискретизации.

Не смотря на достоинства, по сравнению с аналоговыми, цифровые фильтры имеют и ряд недостатков:

- ограничение скорости. Максимальная ширина полосы сигналов, которые в

реальном времени способны обработать цифровые фильтры, значительно уже, чем у аналоговых фильтров, так как ограничивается быстродействием АЦП и ЦАП. Так же, скорость работы цифрового фильтра зависит от скорости работы цифрового процессора и числа арифметических операций, которые необходимо выполнить в алгоритме фильтрации;

- влияние конечной разрядности. Цифровые фильтры подвержены шуму (погрешности) АЦП, вызванного конечным числом уровней квантования (разрядность АЦП), и шуму округления, вызванного конечной разрядностью слова (обычно 8 или 16 бит). При использовании рекурсивных фильтров высоких порядков накопление шума округления может привести к неустойчивости фильтра;

- значительное время разработки и внедрения. Разработка и внедрение цифровых фильтров, особенно внедрение аппаратного обеспечения, могут выполняться гораздо дольше, чем подобные процедуры для аналоговых фильтров. В то же время, однажды разработанное аппаратное и/или программное обеспечение может использоваться в других задачах цифровой обработки сигналов с незначительной модификацией или вообще без изменений.

Пренебрежение влиянием фазочастотной характеристики фильтра на форму сигнала - весьма распространенная ошибка. Фаза важна потому, что сигнал, прошедший через фильтр без изменения амплитуды в полосе пропускания, может быть искажен по форме, если запаздывание при прохождении через фильтр не будет постоянным для разных частот.

Одинаковое время задержки соответствует линейной зависимости фазы от частоты.

Эквалайзеры

На основе отдельных фильтров строятся эквалайзеры (equalizer, EQ), которые объединяют в себе несколько фильтров, предназначенных для изменения спектральных свойств (тембра) обрабатываемого сигнала. Существует два различных по назначению и по устройству типа эквалайзеров: графический и параметрический.

Графический эквалайзер. Это набор полосовых фильтров с фиксированными центральными частотами и переменным коэффициентом усиления, которым можно управлять при помощи слайдера (ползунка). В качестве регуляторов принято использовать именно слайдеры, так как их положение представляет собой некое подобие графика АЧХ эквалайзера. Пользователь как бы рисует слайдерами необходимую ему кривую АЧХ.

На вход группы включенных параллельно фильтров подается один и тот же сигнал, и задача каждого фильтра состоит в том, чтобы усилить или ослабить "свой" участок спектра в соответствии с положением регулятора коэффициента усиления.

Частоты регулирования в графических эквалайзерах унифицированы. Они выбираются из ряда стандартных частот, перекрывающих весь звуковой диапазон и отстоящих друг от друга на некоторый постоянный интервал. Этот интервал может составлять октаву, ее половину, треть. Наиболее широкие возможности, естественно, дают третьоктавные эквалайзеры.

В составе любого звукового редактора профессионального уровня обязательно найдется виртуальный графический эквалайзер.

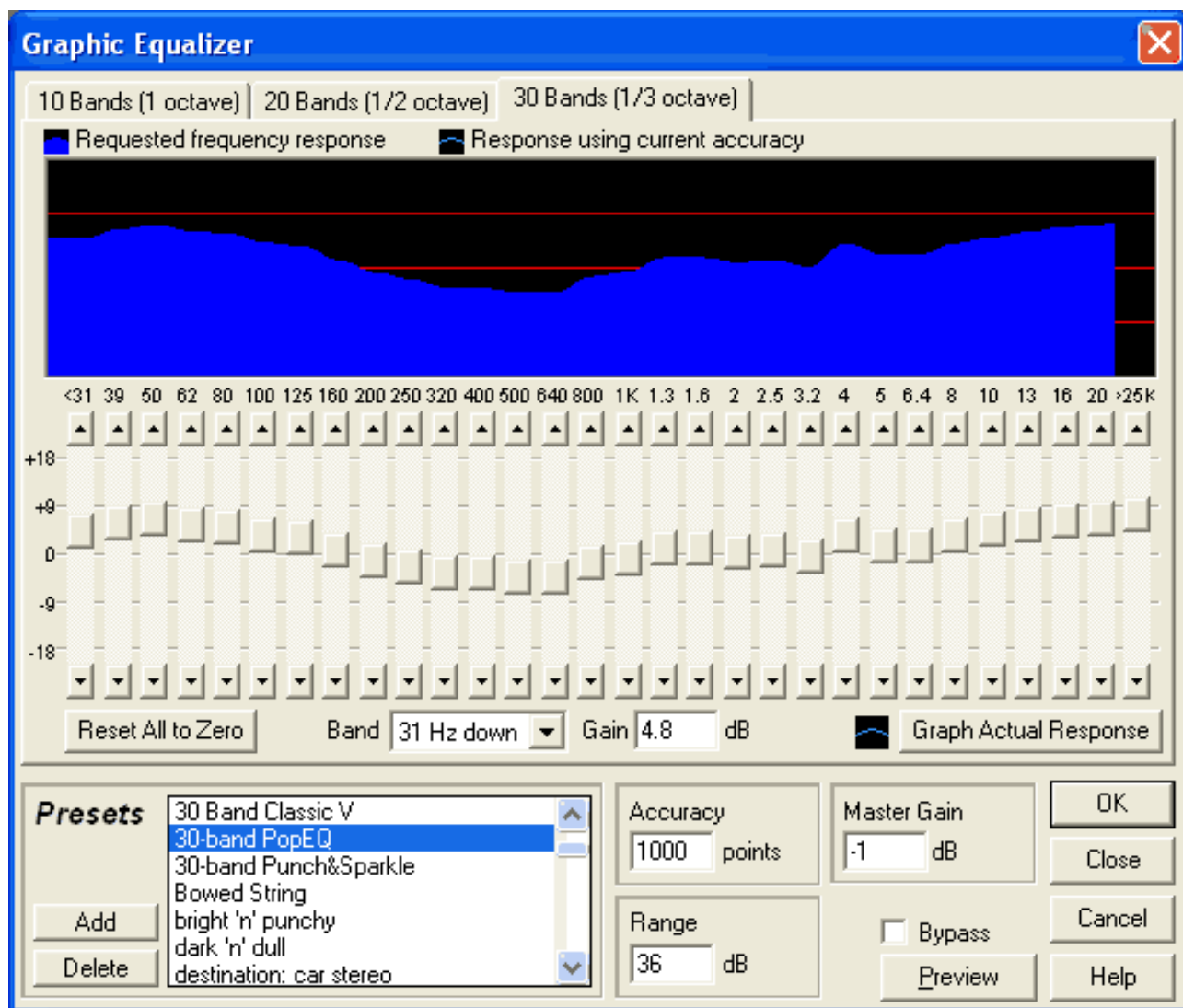


Рис. 5. Графический эквалайзер

Графические эквалайзеры обычно применяются для "доводки" общей картины, обработки суммарного сигнала, а не отдельных составляющих. С помощью графического эквалайзера можно приблизительно сформировать необходимую АЧХ системы обработки звука или акустической системы: поднять усиление в одних областях спектра и уменьшить в других.

Параметрический эквалайзер. Параметрический эквалайзер позволяет управлять не только коэффициентом усиления фильтра, но и его центральной частотой, а также добротностью (по сути дела, шириной полосы пропускания-подавления).

Для формирования АЧХ сложного вида применяются многополосные параметрические эквалайзеры, параметры каждого из них можно изменять независимо. В качестве

примера приведу семиполосный параметрический эквалайзер, реализованный в программе Cool Edit Pro 2. Эквалайзер с АЧХ, представленной на рисунке, позволяет реализовать интересный эффект: из обрабатываемого аудиосигнала выделяются колебания с частотами, которые соответствуют нотам ля пяти соседних октав.

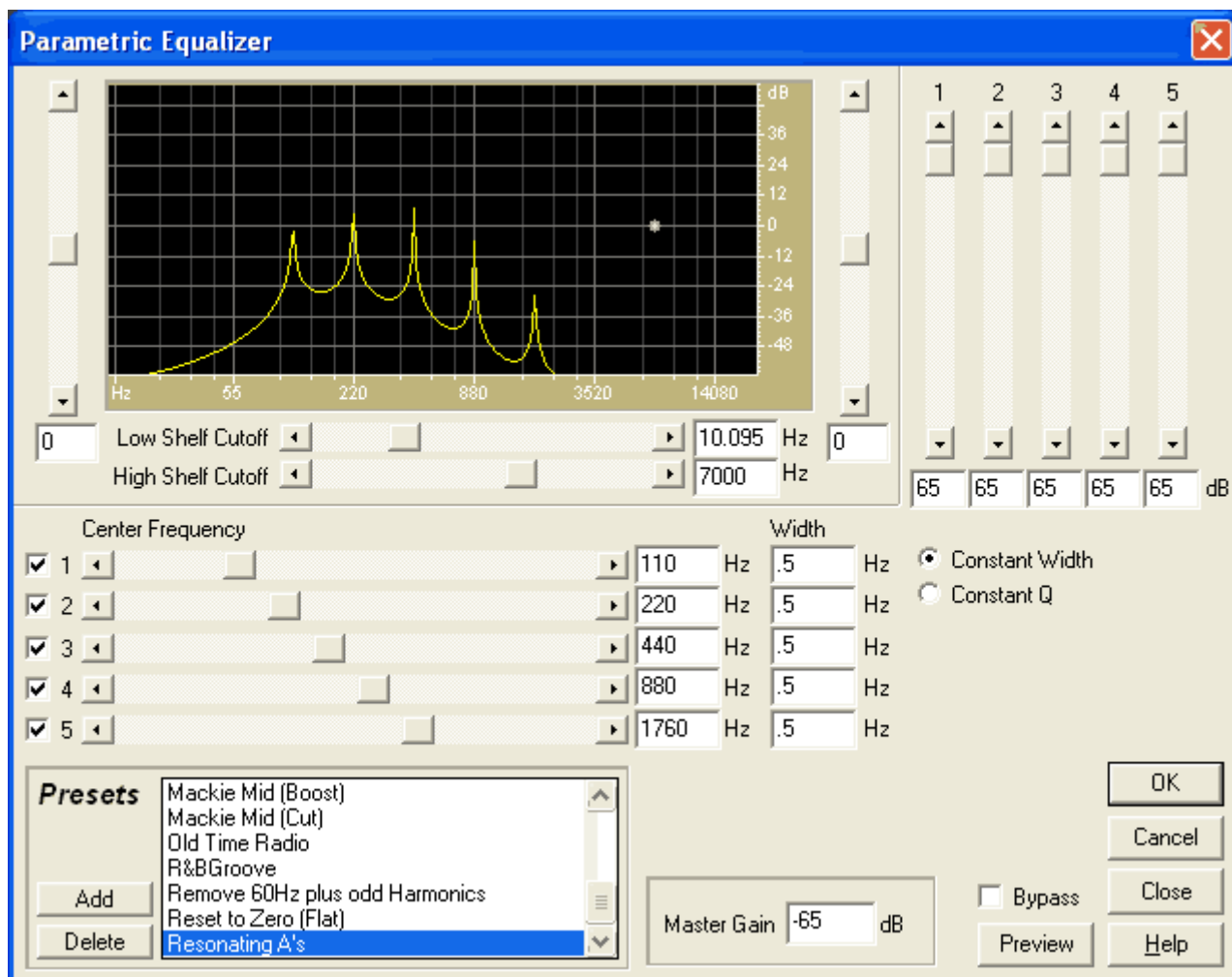


Рис. 6. Параметрический эквалайзер

При наличии некоторого опыта вы сможете точно устанавливать значения параметров эквалайзера таким образом, чтобы подчеркнуть звук отдельного инструмента или удалить нежелательную помеху с минимальным влиянием на остальные элементы звукового образа.

БПФ эквалайзер. В продвинутых современных звуковых редакторах можно задавать любую АЧХ эквалайзера, просто рисуя ее график. Для этого щелчками нужно создать на графике узловые точки, которые затем следует перетаскивать в пределах координатного поля. На рисунке показан фильтр, основанный на быстром преобразовании Фурье (БПФ). Представленная АЧХ позволяет подавить в записанном

сигнале фон электросети (колебание с частотой 50 Гц и ряд его гармоник). Отдельные полосы подавления здесь столь узки, а скаты АЧХ столь круты, что составляющие полезного сигнала будут практически не затронуты, в то время как фон исчезнет полностью.

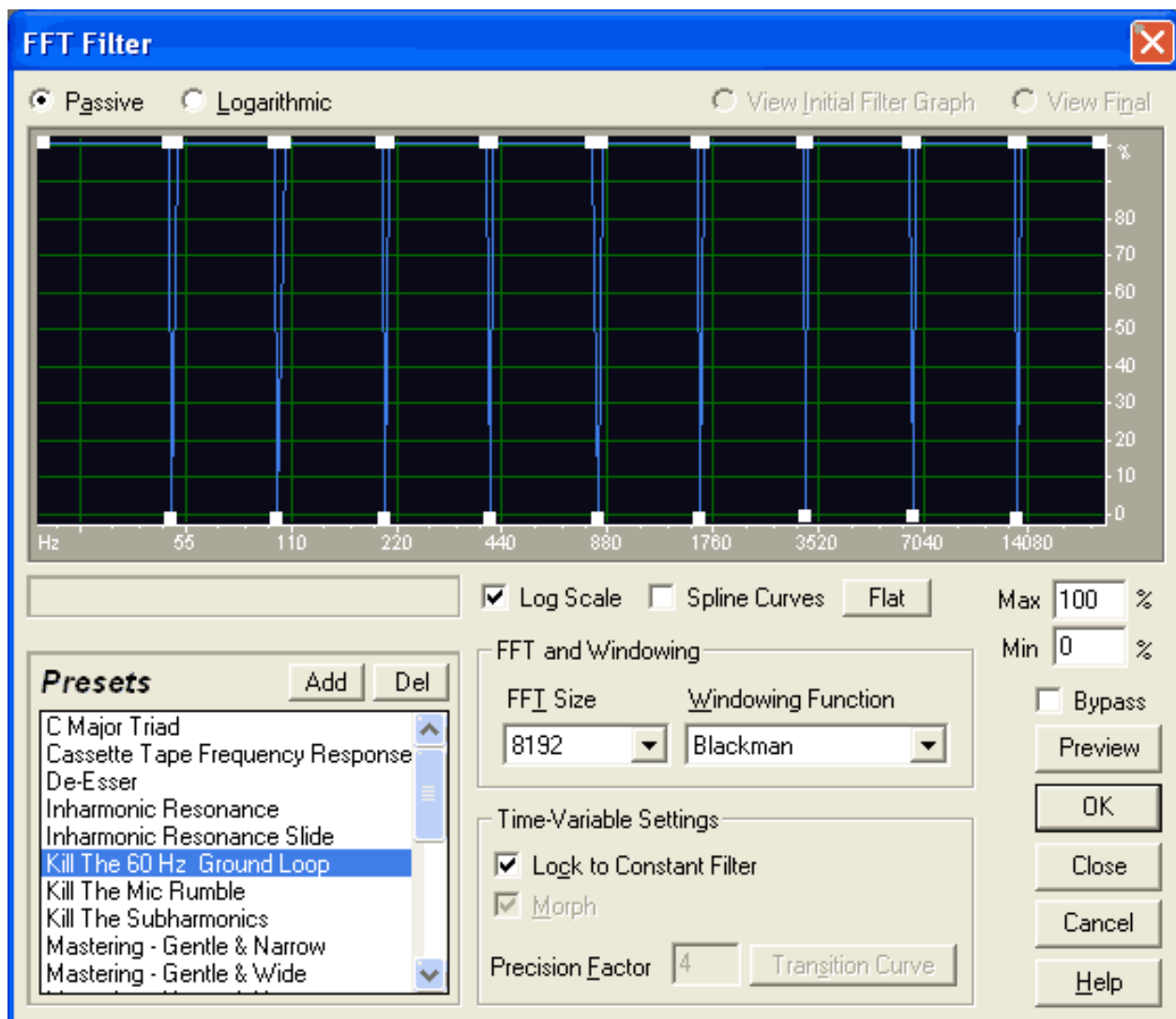


Рис. 7. БПФ- эквалайзер

Эквалайзеры применяют практически на любом этапе любого процесса обработки звука - от записи живого концерта до сведения многоканальной студийной записи. В основном их используют для того, чтобы исправить звуковой сигнал, который не соответствует определенным требованиям.

Лекция 7. Звуковая система ПК

Звуковая система ПК в виде звуковой карты появилась в 1989 г., существенно расширив возможности ПК как технического средства информатизации.

Звуковая система ПК — комплекс программно-аппаратных средств, выполняющих следующие функции:

- ✓ запись звуковых сигналов, поступающих от внешних источников, например, микрофона или магнитофона, путем преобразования входных аналоговых звуковых сигналов в цифровые и последующего сохранения на жестком диске;
- ✓ воспроизведение записанных звуковых данных с помощью внешней акустической системы или головных телефонов (наушников);
- ✓ воспроизведение звуковых компакт-дисков;
- ✓ микширование (смешивание) при записи или воспроизведении сигналов от нескольких источников;
- ✓ одновременная запись и воспроизведение звуковых сигналов (режим *Full Duplex*);
- ✓ обработка звуковых сигналов: редактирование, объединение или разделение фрагментов сигнала, фильтрация, изменение его уровня;
- ✓ обработка звукового сигнала в соответствии с алгоритмами объемного (трехмерного — *3D-Sound*) звучания;
- ✓ генерирование с помощью синтезатора звучания музыкальных Инструментов, а также человеческой речи и других звуков;
- ✓ управление работой внешних электронных музыкальных инструментов через специальный интерфейс MIDI.

Звуковая система ПК конструктивно представляет собой звуковые карты, либо устанавливаемые в слот материнской платы, либо интегрированные на материнскую плату или карту расширения другой подсистемы ПК. Отдельные функциональные модули звуковой системы могут выполняться в виде дочерних плат, устанавливаемых в соответствующие разъемы звуковой карты или подключаться по USB.

Аппаратная часть. Кодеки.

Аппаратная часть типичная звуковая система ПК, как показано на рис.1

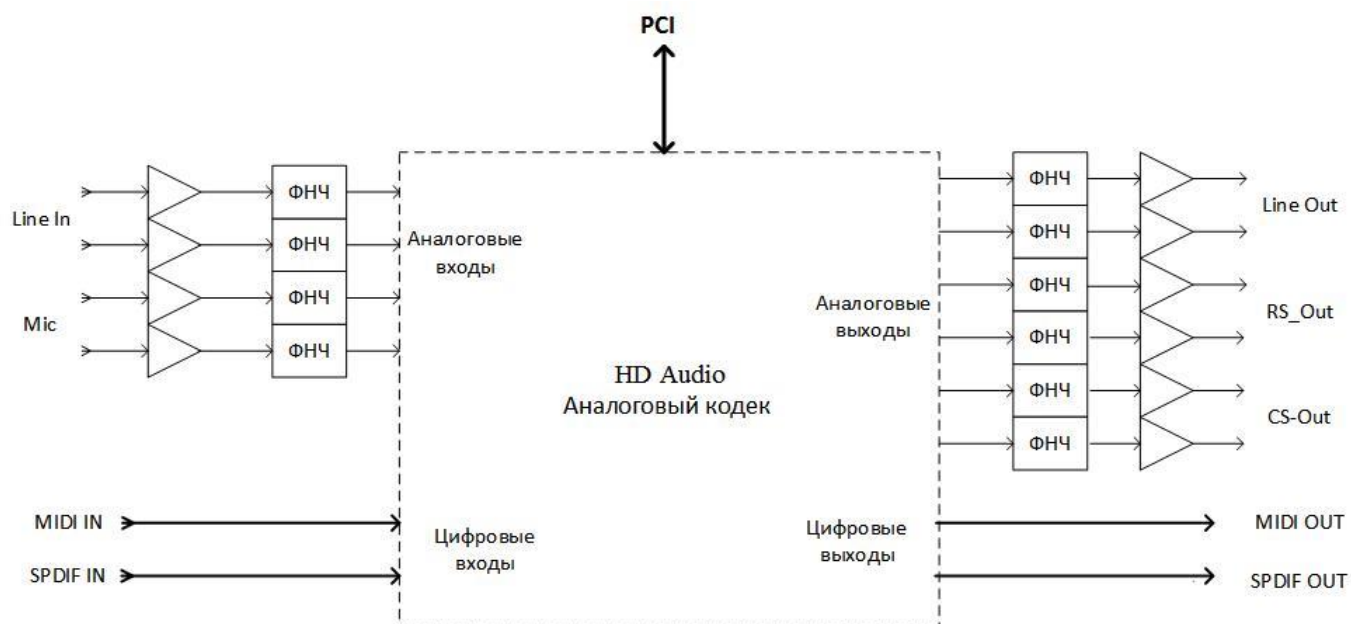


Рисунок 1. Аппаратная часть звуковой системы ПК

Ключевой элемент звуковой системы – аналоговый кодек, который осуществляет аналого-цифровое и цифро-аналоговое преобразования звуковых сигналов, а также реализует прием и передачу звуковых сигналов по цифровым интерфейсам.

AC'97 (сокращенно от англ. *audio codec '97*) — это стандарт для аудиокодеков, разработанный подразделением Intel Architecture Labs компании Intel в 1997 г. AC'97 поддерживает частоту дискретизации 96 кГц при использовании 20-разрядного стереоразрешения и 48 кГц при использовании 20-разрядного стерео для многоканальной записи и воспроизведения.

HD Audio (от англ. *high definition audio* — звук высокой чёткости) является эволюционным продолжением спецификации AC'97, предложенным компанией Intel в 2004 году, обеспечивающим воспроизведение большего количества каналов с более высоким качеством звука, чем при использовании интегрированных аудиокодеков AC'97. Аппаратные средства, основанные на HD Audio, поддерживают 24-разрядное качество звучания (до 192 кГц в стереорежиме, до 96 кГц в многоканальном режимах

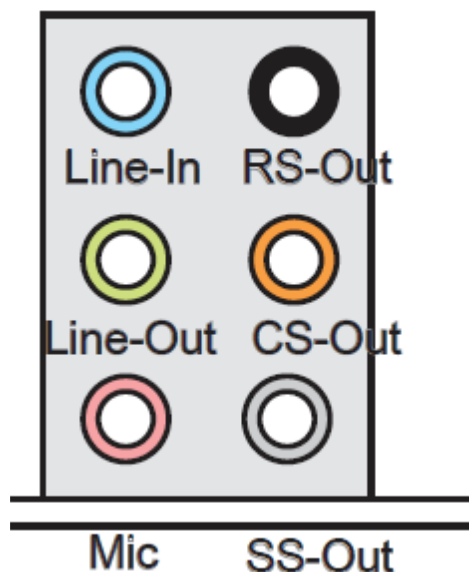
(до 8 каналов)).

Формфактор кодеков и передачи информации между их элементами остался прежним. Изменилось только качество микросхем и подход к обработке звука.

Сравнение спецификаций кодеков

AC '97	HD Audio	Преимущество HD Audio
20 бит 96 кГц максимум	24 бит 192 кГц максимум	Полноценная поддержка новых форматов, таких, как DVD-Audio
2.0	5.1/7.1	Полноценная поддержка новых форматов, таких, как Dolby Digital Surround EX, DTS ES
Полоса пропускания 11,5 Мб/с	48 Мб/с выход, 24 Мб/с вход	Более широкая полоса пропускания позволяет использовать большее число каналов в более детальных форматах
Фиксированная полоса пропускания	Задаваемая полоса пропускания	Используются только необходимые ресурсы
Определённый канал DMA	DMA каналы общего назначения	Поддержка многопоточности и нескольких подобных устройств
Одно звуковое устройство в системе	Несколько логических звуковых устройств	Поддержка концепции Digital Home / Digital Office, вывод разных звуков на разные выходы для мультимедийных возможностей и отдельного голосового чата во время онлайн-игр
Опорная частота задаётся извне, основным кодеком	Опорная частота берётся от чипсета	Единый высококачественный задающий генератор для синхронизации
Стабильность работы зависит от стороннего ПО третьих фирм	Универсальная архитектура звукового драйвера от Microsoft	Единый драйвер для большей стабильности OS и базовой функциональности, не требуется специальная установка драйверов
Ограниченное автоопределение и переопределение	Полное автоопределение и переопределение	Полная поддержка Plug and Play
Стереомикрофон или 2 микрофона	Поддержка массива из 16 микрофонов, максимум	Более точные ввод и распознавание речи

Рассмотрим входы и выходы звуковой система ПК.



Line In (синий) - это **Линейный вход**, позволяющий вводить в звуковые сигналы стандартного уровня (около 250 милливольт или 0,25 Вольт). Это выходы микшера, синтезатора, магнитофона, CD-плеера и пр.

Mic (Розовый) – это **Микрофонный вход** звуковой карты, оснащенный микрофонным предусилителем. На выходе микрофона сигнал очень слабый, поэтому микрофон, подключенный к линейному входу мы не услышим. В некоторых микрофонах иногда встраивают предусилитель.

Line-Out (зеленый) - **Линейный выход**, позволяющий выводить из звуковой карты сигналы стандартного уровня (250 милливольт), используется для подключения наушников или колонки 2/4/6/8-канальном аудио режиме.

RS-Out (черный) - выход для тыльных колонок для 4/6/8-канального аудио режима.

CS-Out (Оранжевый) - выход для центральной колонки и сабвуфера для 6/8-канального аудио режима.

SS-Out (Черный) - выход для боковых колонок для 8-канального аудио режима

SPDIF - это цифровой вход или выход, позволяющий передавать/получать звук на/с записывающее устройство, другую звуковую карту, внешний ЦАП/АЦП, усилитель или ресивер без потерь качества. **SPDIF** или **S/P-DIF** — расшифровывается как Sony/Philips Digital Interface (или Interconnect) Format (описано также как [IEC 958](#) type II в международном стандарте IEC-60958). Является совокупностью спецификаций протокола низкого уровня и аппаратной реализации, описывающих передачу цифрового звука между различными компонентами аудиоаппаратуры. При описании S/PDIF необходимо описать как физическую часть (то есть, собственно, каким образом сигнал передаётся и по чему), так и программную часть (то есть используемый протокол). **S/PDIF** первоначально применялся в CD-плеерах (и DVD-плеерах, проигрывающих компакт-диски), а затем стал общим способом соединения и передачи звука в других аудиокомпонентах. Он также приобрёл популярность в автомобильном звуке, где прежний беспорядок многочисленных проводов может быть заменён единственным волоконно-оптическим кабелем, который устойчив к электрическим помехам.

Другое применение интерфейс **S/PDIF** находит в передаче цифрового потока объёмного звука как определено стандартом IEC 61937. Этот режим используют, чтобы подключить выход DVD-плеера к входу AV-ресивера домашнего кинотеатра, который поддерживает форматы Dolby Digital или Digital Theatre System(DTS) объёмного звука.

Разъём RCA ("тюльпан") — наиболее распространённый разъём, используемый с интерфейсом S/PDIF и идентичный разъёму, применяемому в потребительской аудио продукции.



Кабель и оптический (FIBER) разъем SPDIF



Сейчас большую популярность приобрели разъёмы типа MiniTOSLINK -это разъём

оптического кабеля в форм-факторе 3,5мм jack. Очень часто такие разъёмы встречаются в современных ноутбуках, где выход S/PDIF совмещён с выходом на наушники. Для соединения такого ноутбука с ресивером потребуется кабель MiniTOSLINK — TOSLINK, либо переходник для стандартного кабеля TOSLINK-TOSLINK.



Режимы SPDIF:

- **AES/EBU** - студийный профессиональный формат передачи цифрового звука;
- **PCM** - несжатый звуковой поток;
- **AC-3, DTS, Dolby Digital** - сжатый многоканальный звук. Термин pass-through (транзит) - означает, что звуковая плата без изменения пропускает через себя закодированный (AC-3, DTS, Dolby Digital) многоканальный саундтрек к фильму (в той же форме, в которой он записан на DVD) в ресивер. Декодирование саундтрека производится в ресивере.

Цифровой вход/выход **SPDIF** может использоваться для синхронизации нескольких аудиокарт. Это необходимо для наращивания количества каналов записи и/или воспроизведения.

MIDI вход и выход - это специальные разъёмы для подключения MIDI-клавиатуры или синтезатора:



Обычно подключение ПК в MIDI-сеть осуществляется с помощью специального MIDI-адаптера, который имеет три MIDI-порта: ввода, вывода и сквозной передачи данных, а также два разъема для подключения джойстиков.

Система объёмного звука Dolby Digital

Существует несколько путей создания объёмного звука:

- многоканальная стереофония (в том числе системы Surround Sound)
- бинауральная стереофония

Первый и самый простейший метод — это использование микрофонных систем для пространственной звукозаписи и/или сведение объёмного звука для систем громкоговорителей, окружающих слушателя при воспроизведении звука с разных сторон.

Вторая технология — преобразование звука с учётом психоакустических методов локализации звука для моделирования двухмерного звукового поля при помощи наушников.

Dolby Digital (AC-3, ATSC A/52) - система пространственного звуковоспроизведения, разработанная фирмой "Dolby Laboratories, Inc." ("Dolby Labs"), руководителем которой является Рей Долби. Впервые технология **Dolby Digital** была продемонстрирована зрителям в июне 1992 года в фильме “Бэтмен возвращается” (“Batman Returns”) и с тех пор вышла уже не одна тысяча лент со звуком в этом формате. Более того, в настоящее время формат **Dolby Digital** в США принят в качестве звукового стандарта для телевидения высокой четкости (HDTV), используется для передачи по спутниковым и кабельным каналам.

Dolby Digital предоставляет в общей сложности **шесть** отдельных каналов звука. Она включает в себя левый, центральный и правый каналы во фронтальной части комнаты, левый и правый тыловые каналы объёмного звука для более точного определения местоположения звуков и более натуральной, реалистичной передачи атмосферы и фона. Пять основных каналов передают полный спектр частот (от 3 до 20000 Гц). Шестой канал - Low Frequency Effects Channel (канал для низкой частоты и эффектов) сильно ограничен сверху по частоте (от 3 до 120Гц), его иногда называют ".1" каналом. К этому каналу подключают низкочастотные динамики (сабвуферы). Если его добавляют к полным 5 каналам Dolby Digital, то про такие системы говорят, как про

имеющие "5.1" канала. Все 6 каналов звука закодированы в один файл, который распаковывается с помощью специального декодера и разводится на шесть аудиоколонок.

Dolby Digital Plus – использует 8 каналов ("7.1"). Конфигурации и идентификация каналов приведена ниже.

Конфигурации каналов звука (громкоговорителей)

Фронтальный левый	Центральный	Фронтальный правый
Окружной левый		Окружной правый
Окружной тыловой левый		Окружной тыловой правый
Низкочастотный		

Параметры оценки качества звукового тракта

Для оценки качества звукового тракта используются следующие параметры:

- диапазон рабочих частот;
- коэффициент нелинейных искажений;
- отношение сигнал/шум.
- динамический диапазон

Диапазон рабочих частот это диапазон звуковых частот, в котором неравномерность АЧХ относительно средней частоты 1000 Гц, как правило, не превышает 6 дБ. АЧХ показывает, насколько равномерно может воспроизводить частоты звуковая система. Считается, что человеческое ухо воспринимает частоты в интервале 20 Гц...20 кГц. И если аудиосистема воспроизводит частоты в том же диапазоне, то ее звучание будет наиболее естественным. Важен не только широкий диапазон воспроизведения частот, но и вид кривой АЧХ. Хорошо, если АЧХ представляет собой прямую линию в пределах достаточно широкого диапазона воспроизводимых частот. Для большинства звуковых плат этот диапазон составляет от 30 Гц до 20 кГц.

Коэффициент нелинейных искажений (total harmonic distortion (THD)). Если на вход звукового устройства подается идеальный синусоидальный сигнал (основная гармоника), то на выходе его форма уже будет отлична от первоначальной. Вот это искажение сигнала и характеризует *THD*.

$$THD = \frac{\sqrt{V_2^2 + V_3^2 + V_4^2 + V_5^2}}{V_1}$$

где V_i -среднеквадратическое (действующее) значение (СКО) напряжения i -й гармоники. V_1 - СКО основной гармоники.

Каждый нелинейный элемент является причиной искажения. Чем меньше этот коэффициент, тем выше качество звука. Если эта величина превышает 0,1 %, то помехи могут быть заметны на слух. Устройство с таким уровнем помех не может считаться высококачественным. Для непрофессиональных звуковых карт THD в

пределах 0,004...0,02% можно считать приемлемым результатом.

Отношение сигнал/шум (signal-to-noise ratio SNR). Вместе с воспроизводимым сигналом на выходе звукового устройства всегда присутствует шум. Чем меньше его амплитуда по отношению к основному сигналу, тем лучше. Особенно хорошо шум проявляется при прослушивании музыки в наушниках и не так заметен при удалении от акустических колонок. Отношение сигнал/шум вычисляется по формуле

$$SNR = 20 \log_{10} \frac{S}{N}$$

где N -среднеквадратическое (действующее) значение (СКО) напряжения шума;

S - среднеквадратическое (действующее) значение (СКО) напряжения сигнала.

Динамический диапазон (dynamic range DR). Динамический диапазон устройства преобразования и обработки звука может быть определен выражением

$$DR = 20 \log_{10} \frac{S_{max}}{S_{min}}$$

где S_{max} и S_{min} — максимальное и минимальное значения сигнала.

Увеличение разрядности АЦП обусловлено еще одним фактором — стремлением расширить его динамический диапазон. Очевидно, минимальный сигнал не может быть меньше, чем напряжение, соответствующее одному разряду, а максимальный — не должен превышать величины напряжения, соответствующего N разрядам (битам).

Динамический диапазон N - битного АЦП:

$$D \leq 6N + 1,76 \text{ дБ.}$$

Иными словами, для записи звучания некоторого источника звука, динамический диапазон которого составляет 120 дБ, требуется двадцатиразрядный АЦП. Если такого нет, а имеется только шестнадцатиразрядный, то динамический диапазон звука должен быть сжат на 24 дБ: со 120 дБ до 96 дБ.

Так, естественный динамический диапазон симфонического оркестра определяется значением максимального звукового давления (110 дБ — самые громкие музыкальные звуки при игре всех инструментов) и значением минимального

звукового давления (30дБ — тихая игра, например, одного инструмента при исполнении сольной партии).

Также стоит определиться с единицами, в которых обычно представлены параметры звука. Для этого используется относительная логарифмическая единица — децибел (дБ, dB). В технических описаниях можно встретить некоторый разнобой. Скажем, в одних документах значение SNR указывается со знаком «—», в других — без такового. В общем-то, сути это не меняет, нужно обращать внимание на модуль числа этого параметра. Знак «—» перед дБ обозначает ослабление, его указание было бы правильнее, но, по-видимому, считается не обязательным. Уровень звука также измеряется в дБ и тоже записывается со знаком «—». Максимальный уровень звука конкретного устройства принято считать равным 0 дБ, другие значения записываются с отрицательным знаком.

Лекция 8. Кодирование звуковой информации

Для хранения цифрового звука существует много различных способов. Оцифрованный звук представляет собой набор значений амплитуды сигнала, взятых через определенные промежутки времени.

Прямой, обратный, дополнительный коды

Целые числа хранятся в компьютере в двоичном формате. Неотрицательные и отрицательные числа хранятся в памяти компьютера по-разному. Один, старший, двоичный разряд отводится под обозначение знака числа. Ноль в старшем разряде означает, что хранится неотрицательное число, единица означает, что число отрицательное. Применяются три формы кодирования целых чисел: прямой код, обратный код, дополнительный код.

Прямой код. Для представления числа в прямом коде n -разрядного формата нужно перевести число в двоичную систему счисления и дополнить слева нулями до n знаков. Так как старший разряд числа отводится для знака, а оставшиеся $n - 1$ разрядов - для значащих цифр, то в знаковый разряд записать 1, если число отрицательное, и оставить 0, если число положительное.

Например, формат хранения целого однобайтного числа имеет вид:



Таким образом, число 3_{10} в прямом коде однобайтного формата будет представлено в виде:

0	0	0	0	0	0	1	1
---	---	---	---	---	---	---	---

Отрицательное число -3_{10} в прямом коде однобайтного формата имеет вид:

1	0	0	0	0	0	1	1
---	---	---	---	---	---	---	---

Использование чисел со знаком (прямого кода представления чисел) усложняет структуру компьютера. В этом случае операция сложения двух чисел, имеющих разные знаки, должна быть заменена на операцию вычитания меньшей величины из большей и присвоения результату знака большей величины. Поэтому в современных компьютерах, как правило, отрицательные числа представляют в виде дополнительного или обратного кодов, что при суммировании двух чисел с разными знаками позволяет заменить вычитание на обычное сложение.

Обратный код. Для представления отрицательного числа в обратном коде n -разрядного формата нужно модуль отрицательного числа записать в прямом коде n двоичных разрядах (перевести число в двоичную систему счисления и дополнить слева нулями до n бит). Значения всех бит инвертировать (нули заменить единицами, единицы нулями)

Дополнительный код. Для представления отрицательного числа в дополнительном коде n -разрядного формата нужно представить его в обратном коде и прибавить 1 к последнему биту числа.

Заметим, что положительные целые числа в прямом, обратном и дополнительном кодах изображаются одинаково - двоичными кодами с цифрой 0 в знаковом разряде.

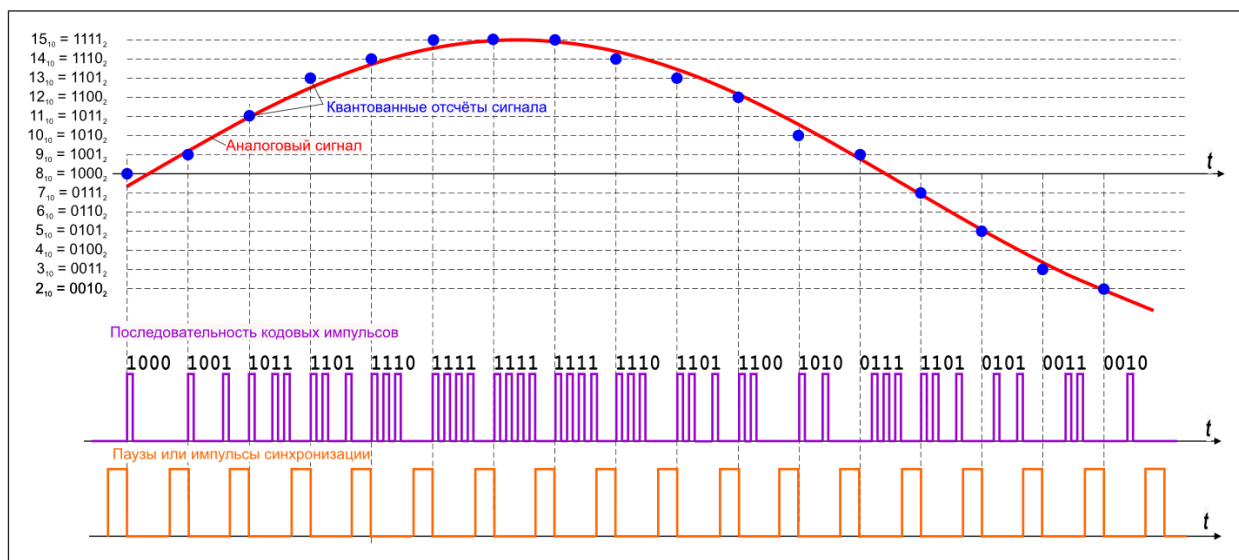
Импульсно-кодовая модуляция

Отведём для записи одного значения амплитуды сигнала в памяти компьютера N бит. Значит, с помощью одного N -битного слова можно описать 2^N разных значений.

Квантование по амплитуде — процесс замены реальных значений амплитуды сигнала значениями, приближенными с некоторой точностью. Каждый из 2^N возможных уровней называется уровнем квантования, а расстояние между двумя ближайшими уровнями квантования называется шагом квантования. Если амплитудная шкала разбита на уровни *линейно*, то шаг квантования равен $\Delta = \frac{1}{2^N}$. Этот способ оцифровки сигнала — дискретизация сигнала во времени в совокупности с методом линейного квантования — называется линейной импульсно-кодовой модуляцией (ИКМ) (Linear Pulse Code Modulation — LPCM) или просто ИКМ (PCM).

Для записи звука на цифровые носители или передачи по каналам связи (см.

интерфейс SPDIF) данные в параллельном коде подвергаются кодированию при помощи регистра сдвига, тактируемого вспомогательным генератором. На выходе регистра сдвига формируются пакеты кодированных импульсов в последовательном коде.



Стандартный аудио компакт-диск (CD-DA), применяющийся с начала 80-х годов 20-го столетия, хранит информацию в формате РСМ с частотой дискретизации 44.1 кГц и разрядностью квантования 16 бит.

Существуют разновидности ИКМ, в которых используется неравномерный шаг квантования (Nonuniform PCM). Способ неравномерного (нелинейного) *квантования* предусматривает разбиение амплитудной шкалы на уровни, по нелинейному (как правило, по логарифмическому) закону. Такой способ квантования называют *логарифмическим квантованием*. При использовании логарифмической амплитудной шкалы, в области слабой амплитуды оказывается большее число уровней квантования, чем в области сильной амплитуды (при этом, общее число уровней квантования остается таким же, как и в случае однородного квантования). *Относительная погрешность – const.*

Альтернативным способом аналогово-цифрового преобразования является дифференциальная (разностная) *импульсно-кодовая модуляция* ДИКМ (Differential PCM). В случае ДИКМ квантованию подвергают не саму амплитуду, а разность между

текущим и предыдущим измерениями. В полной аналогии с ИКМ, разностная ИКМ может сочетаться с использованием как равномерного, так и неравномерного квантования. Для звуковых данных такой тип модуляции уменьшает требуемое количество бит на отсчёт примерно на 25%.

Разностное кодирование имеет много разных вариантов. Например, адаптивная ДИКМ (АДИКМ, ADPCM) — разновидность ДИКМ с переменным шагом квантования. Изменение шага позволяет уменьшить требования к полосе пропускания при заданном соотношении сигнала и шума.

Блок оцифрованной аудио информации можно записать в файл без изменений, то есть последовательностью значений амплитуды, описывающих форму аудио сигнала, т.е. в несжатом виде. Для этого обычно используется формат **.WAV**. *Waveform Audio File Format (WAVE, WAV, от waveform)* — формат файла-контейнера для хранения записи оцифрованного аудио потока, подвид RIFF. Этот контейнер, как правило, используется для хранения несжатого звука в импульсно-кодовой модуляции. Однако контейнер не налагает каких-либо ограничений на используемый алгоритм кодирования

RIFF (*Resource Interchange File Format*) — один из форматов файлов-контейнеров для хранения потоковых мультимедиа-данных (видео, аудио, возможно текст). Наиболее известными форматами, использующими RIFF в качестве контейнера, являются: AVI (видео), WAV (аудио), RMI (MIDI-треки).

Сжатие данных

Как мы выяснили, путем простой оцифровки реального звукового сигнала с высокой частотой выборки и разрядностью квантования можно очень качественно, почти без потерь качества, сохранить желаемый аудио материал в компьютере. Тогда возникает вопрос: зачем нужны различные алгоритмы сжатия звукового материала, подобные MP3? Согласно теореме Котельникова (Найквиста) частота дискретизации устанавливает верхнюю границу частот оцифрованного сигнала, а именно: максимальная частота спектральных составляющих оцифрованного сигнала равна половине частоты дискретизации сигнала - т.е., чтобы получить полную информацию об оригинальном аналоговом сигнале в частотной полосе от 0 до 22050 Гц (в слышимом человеком диапазоне частот), необходимо дискретизовать сигнал как минимум с частотой не менее 44.1 КГц. Это означает, что чтобы сохранить оригинальное качество аудио материала необходимо выбирать высокие значения параметров оцифровки. Однако, чем выше значения этих параметров, тем больший объем памяти компьютера требуется для хранения оцифрованных данных. Например, стандартный аудио компакт диск объемом **650 Мб** хранит аудио данные в формате РСМ 44.1 КГц / 16 бит/stereo. Такие параметры соответствуют двухканальной записи с 65536 (2^{16}) уровнями квантования амплитуды при частоте оцифровки 44100 Гц. Несложно подсчитать, что в таком виде на диск помещается **около часа** музыки, что, в принципе, не очень много, если учесть, что коллекция аудио может исчисляться тысячами часов.

Таким образом, чтобы иметь возможность хранить относительно большие объемы аудио данных в хорошем качестве приходится прибегать к «ухищрениям», которые помогают записать аудио данные с использованием ощутимо меньшего объема памяти (то есть, уплотнить, сжать, закодировать данные) и не слишком сильно ухудшая (или даже совсем не ухудшая) при этом качество звучания. Об этих ухищрениях мы сейчас и поговорим.

Существует два распространенных способа кодирования аудио информации (кроме простого хранения в оцифрованном виде «как есть»).

Кодирование информации – представление информации в определенной системе

кодовых символов и их структур. Шифрование, а также уплотнение (сжатие) информации являются частными случаями кодирования.

1. Сжатие данных без потерь (*lossless coding*) - это способ кодирования аудио информации, который позволяет осуществлять стопроцентное восстановление данных из сжатого потока. К такому способу уплотнения данных прибегают в тех случаях, когда необходимо именно стопроцентное сохранение оригинального качества звучания аудио данных. Например, после сведения звука в студии звукозаписи, данные необходимо сохранить в архиве в оригинальном качестве для их возможного использования впоследствии. Существующие сегодня алгоритмы сжатия без потерь (например, алгоритм, реализованный в кодеке Monkeys Audio , а также кодеках Flac , WavPack , TTA , OptimFrog и других) позволяют сократить занимаемый данными объем на 20-50% и при этом обеспечить стопроцентное восстановление оригинальных данных из полученных после сжатия. Подобные кодеры – это своего рода архиваторы данных (как, например, ZIP, RAR и другие), только предназначенные специально для сжатия аудио информации.

Кодер – программа (или устройство), реализующая определенный алгоритм кодирования данных (например, архиватор, или кодер MP 3), которая в качестве ввода принимает исходную информацию, а в качестве вывода возвращает закодированную информацию в определенном формате.

Декодер – программа (или устройство), реализующая обратное преобразование закодированного сигнала в декодированный.

Кодек (от англ. «codec» - «Coder / Decoder») - программный или аппаратный блок, предназначенный для кодирования/декодирования данных.

Такой способ сжатия данных, хотя и идеален с точки зрения сохранения качества аудио материалов, но неспособен обеспечить высокий уровень компрессии.

2. Сжатие данных с потерями (*lossy coding*). Цель такого кодирования - любыми способами добиться схожести звучания декодированного аудио сигнала с оригиналом при как можно меньшем объеме упакованных данных. Сегодня эта цель достигается за счет использования различных алгоритмов «упрощающих» оригинальный сигнал, выкидывая из него «ненужные» слабослышимые (или вообще неразличимые

человеческим ухом) детали. После такого кодирования, декодированный сигнал при воспроизведении звучит похоже на оригинал, но фактически перестает быть ему идентичным. Методов сжатия, а также программ, реализующих эти методы, существует много. Наиболее известными являются MPEG-1 Layer 3 (это и есть официальное название всем известного «MP3»), MPEG-2/4 AAC (MPEG-2 и MPEG-4 Advanced Audio Coding), Ogg Vorbis (сокращенно OGG), Windows Media Audio (WMA), MusePack (MPC) и другие. Выигрыш от использования таких алгоритмов сжатия вполне очевиден: коэффициент сжатия, обеспечиваемый такими кодерами, находится, усреднено, в пределах 7-14 (раз) и это при малозаметных потерях качества оригинального звучания. Практически это означает, что если одна песня с компакт диска занимает в формате PCM 44.1 КГц / 16 бит / stereo около 50 Мб (в виде .WAV-файла), то она же, сжатая в MPEG-1 Layer 3 (MP3), будет занимать уже около 3-7 Мб и при этом более чем удовлетворять по качеству звучания. Согласитесь, это неплохой вариант.

Как мы сказали, компрессия данных в lossy-кодерах достигается за счет «упрощения» аудио информации. В основе почти всех lossy -кодеров лежит использование так называемой психоакустической модели, которая как раз и используется для упрощения оригинального сигнала. Механизм кодера, основанного на упрощении спектра входного сигнала (есть также кодеры, основанные на других методах) работает приблизительно так. Кодер выполняет анализ кодируемого сигнала, в процессе которого определяются участки сигнала, в определенных частотных областях которых имеются неслышимые человеческому уху нюансы (замаскированные или слабо слышимые частоты, кратковременные малозаметные всплески частот и проч.), после чего происходит удаление подобных нюансов из оригинального сигнала. Такая обработка упрощает форму оригинальной звуковой волны, делая ее «более гладкой». Степень сжатия оригинального сигнала зависит от степени его «упрощения»; сильное сжатие достигается путем «агрессивного упрощения», когда кодер «считает» ненужными множественные нюансы оригинальной звуковой волны. Такое сжатие, естественно, приводит к сильной деградации качества, поскольку удалению могут подлежать не только незаметные, но и значимые детали звучания. Отличительной особенностью всех современных lossy-кодеров является возможность тонкой настройки процесса

кодирования, что, при правильном понимании и подходе, позволяет добиваться высокого коэффициента сжатия при совершенно незаметных (даже на самой хорошей аудио аппаратуре) потерях качества звучания.

Теперь о том, каким образом происходит упрощение сигнала в процессе кодирования (опять же, на примере механизма работы кодера, основанного на упрощении спектра обрабатываемого сигнала). Механизм «упрощения» аудио сигнала можно пояснить следующим образом. Исходный аудио сигнал разделяется на блоки определенной длины, после чего каждый блок обрабатывается в отдельности. В процессе кодирования каждый блок раскладывается на частотные составляющие спектра. Как мы сказали, чем меньше «тонких нюансов звучания» имеется в сигнале (чем проще частотный спектр сигнала), тем эффективнее его компрессия. Упростить сигнал можно по-разному. Например, можно отфильтровать все частотные составляющие сигнала, располагающиеся выше определенной границы, что автоматически упростит сигнал в высоких частотных областях (но при этом, естественно, заметно испортит звучание). Однако основным способом является проведение психоакустического анализа с последующей соответствующей обработкой сигнала: кодер анализирует аудио информацию и, опираясь на указанный пользователем битрейт, «решает» какие тонкости звучания можно выбросить.

Битрейт – количество бит, используемых для хранения одной секунды аудио. Для стандартного . WAV -файла в формате PCM 44.1 KHz / 16 bit / stereo это число составляет: 44100 (значений амплитуды в секунду) * 16 (бит для записи одного значения амплитуды) * 2 (канала) = 1411200 бит в секунду = около 1378 Кбит/с (килобит в секунду).

При сжатии в MP 3 (и многие другие форматы) пользователь указывает желаемый битрейт (или границы изменения битрейта) для сжатого выходного потока. Чем ниже битрейт, тем меньше бит позволяет кодеру отводить для хранения информации об одной секунде аудио и, таким образом, тем сильнее кодер упрощает сигнал, что соответственно влияет на качество звучания получаемого сжатого потока аудио. Наиболее распространенное в Интернете среднее значение битрейта для музыки в формате MP 3 колеблется в пределах от 128 до 192 Кбит/с.

Следует особо подчеркнуть, что применение психоакустики в качестве механизма упрощения приводит к тому, что процесс декодирования уже не способен восстановить утраченные во время компрессии данные (выброшенные нюансы звучания, отфильтрованные частоты и проч.). Это означает, что сжимать данные нужно «с умом»: если вы сжимаете музыку для вашей аудио коллекции, то не следует слишком ограничивать битрейт кодера, поскольку это может привести к серьезным потерям качества. С другой стороны, если подходить к кодированию компетентно, то результатом кодирования будет хороший коэффициент сжатия (что и является целью кодирования) плюс высокое качество звучания.

Лекция 9. Синтез музыки

Интерфейс MIDI

MIDI (англ. *Musical Instrument Digital Interface* — цифровой интерфейс музыкальных инструментов) — стандарт цифровой звукозаписи на формат обмена данными между электронными музыкальными инструментами.

Интерфейс позволяет единообразно кодировать в цифровой форме такие данные как нажатие клавиш, настройку громкости и других акустических параметров, выбор тембра, темпа, тональности и др., с точной привязкой во времени. В системе кодировок присутствует множество свободных команд, которые производители, программисты и пользователи могут использовать по своему усмотрению. Поэтому интерфейс MIDI позволяет, помимо исполнения музыки, синхронизировать управление другим оборудованием, например, осветительным, пиротехническим и т. п.

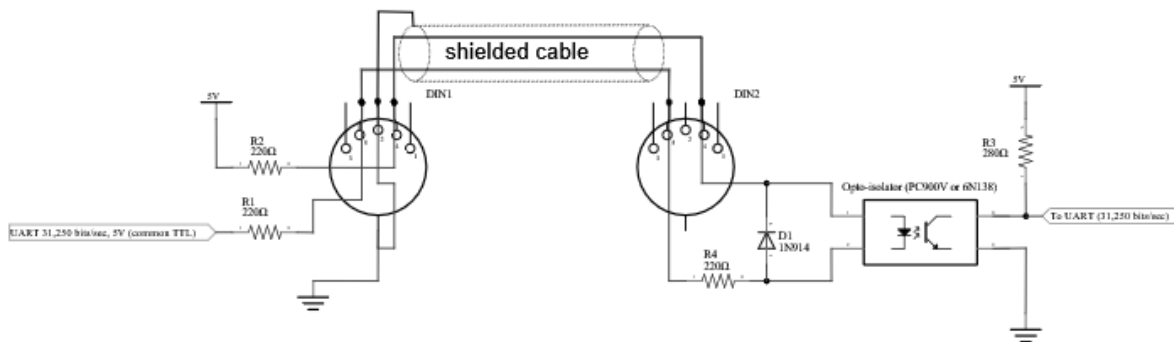
Последовательность MIDI-команд может быть записана на любой цифровой носитель в виде файла, передана по любым каналам связи. Воспроизводящее устройство или программа называется *синтезатором (секвенсором) MIDI* и фактически является автоматическим музыкальным инструментом.

К концу 70-х годов XX века получили распространение музыкальные синтезаторы. Они представляли собой наборы генераторов звуковых частот, управляемых напряжением. Нажатие клавиши на клавиатуре синтезатора включало генератор, частота которого задавалась напряжением от регулятора данной клавиши.

Синтезаторы в то время были полностью аналоговые, все внутренние их блоки (звуковые генераторы, генераторы огибающей, фильтры) управлялись напряжением. Например, звуковой генератор инструмента при подаче напряжения в 1В мог давать высоту тона 100 Гц, 2 В — 200 Гц, 3 В — 400 Гц и так далее. Очевидно, что для внешнего управления таким прибором мог использоваться только аналоговый интерфейс. Он имел название CV/Gate. На вход CV подавалось управляющее напряжение (Control Voltage), пропорциональное высоте ноты, на вход Gate — импульс (trigger), от которого стартовала и выключалась нота.

Количество отдельных генераторов определяло количество тонов синтезатора, которые могут звучать одновременно. Конкретные модели синтезаторов могли иметь особенности звучания и характерные специальные эффекты. Реализовать управление всей гаммой потенциальных возможностей синтеза звука в рамках одного аналогового устройства было невозможно. Рабочее место музыканта того времени могло состоять из нескольких разнородных синтезаторов; работать с таким зоопарком было сложно. В начале 80-х годов необходимость увеличения гибкости управления разнородными синтезаторами стала очевидной, а развитие электроники подсказало путь решения проблемы: цифровое программное управление. Компании-производители синтезаторов удачно смогли договориться о разработке и поддержке единого стандарта на интерфейс управления синтезаторами, который и появился в 1982 году. Идеология стандарта подразумевала разделение органов управления и аппаратуры синтеза звука. Интерфейс MIDI фактически стал средством передачи положения клавиш и регуляторов от музыкальной клавиатуры к аппаратуре синтеза звука. Теперь музыкант мог с одной клавиатуры управлять несколькими синтезаторами разных типов и фирм-изготовителей. Более того, цифровой формат передачи позволил дополнить аппаратуру секвенсорами — устройствами, запоминающими наигранную мелодию. Теперь музыкант мог записывать композиции или их части, микшировать без потери качества звука и использования сложных студийных магнитофонов того времени. Стандарт быстро завоевал популярность. Постепенно дополняясь новыми возможностями, он не потерял актуальности и в наше время.

Физический уровень интерфейса представляет собой токовую петлю (что обеспечивает гальваническую развязку и безопасность при соединении устройств между собой). Передатчик активный, 0..5 мА, наличие тока обозначает 0, нет тока — 1. Разъем 5-штырьковый DIN 41524. Приёмопередатчик асинхронный, скорость 31,25 кбит/с, формат 8-N-1.



MIDI имеет шинную топологию. Источник данных (например, MIDI клавиатура) имеет разъем выхода данных (MIDI OUT), который соединяется кабелем с входом (MIDI IN) первого синтезатора. Синтезатор имеет второй разъем (MIDI THRU), на котором дублируется поток данных со входа. Это позволяет соединить в цепочку произвольное количество синтезаторов. Синтезатор с клавиатурой имеет все три разъема - IN, OUT, THRU.

Интерфейс однонаправленный (симплексный), то есть источник сообщений (например, музыкальная клавиатура) только передает. Никакой обратной связи не предусмотрено.

Стандарт описывает аппаратный интерфейс, который позволяет соединять электронные музыкальные инструменты и компьютеры различных производителей, описывает протоколы связи для передачи данных от одного устройства к другому. MIDI-устройства могут взаимодействовать с программными приложениями, используя коммуникационный протокол MIDI. Используя соответствующий программный MIDI-секвенсор, внешние MIDI-устройства могут посылать информацию на синтезатор звуковой карты. MIDI базируется на пакетах данных, каждый из которых соответствует MIDI-событию (англ. *MIDI-events*), от нажатия клавиши до простой паузы, эти события разделяются по каналам. Сложная среда MIDI может включать различную аппаратуру, причём каждая часть системы будет отвечать за события на соответствующем канале. Альтернативным вариантом может быть одиночный синтезатор, который сам может управлять всеми каналами.

MIDI — это протокол связи между устройством управления, генерирующим команды, и подчиненным устройством, выполняющим эти команды. Если очень сильно сузить

это определение, то можно привести типичный пример: MIDI позволяет исполнителю нажать клавишу на одном инструменте, а получить при этом звук другого или даже нескольких. Любые воздействия исполнителя на органы управления (нажатие клавиш, педалей, изменение положений регуляторов и т. п.) могут быть преобразованы в команды, которые можно передать по MIDI-кабелю на другие инструменты. Эти инструменты, получая команды, обрабатывают их так же, как и при воздействии на их собственные органы управления.

В качестве мастер-устройства может выступать любой прибор, имеющий MIDI-выход (MIDI Out) и способный посылать на этот выход команды управления. Мастер-устройства можно разделить на два типа: устройства, на которые непосредственно воздействует исполнитель (например, синтезатор) и устройства, которые генерируют управляющие команды автоматически (без участия исполнителя), на основе ранее введенных данных. Типичным примером устройства последнего типа является секвенсор.

Протокол MIDI разрабатывался для управления синтезаторами, а в них, как известно, самый главный орган управления — клавиатура. Неудивительно поэтому, что разработчики MIDI для описания действий исполнителя выбрали принцип клавишного инструмента.

MIDI является выраженным клавишно-ориентированным протоколом.

Это не означает, что управлять тон-генератором можно только с клавиатуры — существуют множество других способов ввода, например, электронные пэды и целые ударные установки, гитарные или духовые контроллеры (о них мы поговорим отдельно и более подробно). Однако, какое бы средство ввода не использовалось, сообщения от него преобразуются в клавишно-ориентированные.

Приемы звукоизвлечения, нехарактерные для клавишного инструмента, могут быть лишь симитированы средствами MIDI с той или иной степенью достоверности.

MIDI — протокол реального времени Это означает, что вся система работает по

принципу "получил — выполнил". Исполнитель нажимает клавишу, клавиатура генерирует сообщение "взять ноту" и передает его на вход тон-генератора. Тон-генератор немедленно воспроизводит ноту. Таким образом, никакие параметры, связанные с моментами выполнения команд, в сообщениях не передаются. Моментом выполнения считается момент получения команды. Поэтому сообщения вроде "сыграть ноту через две секунды" в MIDI отсутствуют. Как это ни странно, сообщение "сыграть ноту длительностью две секунды" в системе реального времени также невозможно.

Компоненты MIDI Понятия, затронутые выше, также требуют подробного обсуждения, но уже сейчас ясно, из чего должен состоять протокол. Очевидно, для того, чтобы одно устройство реагировало на управляющие воздействия с другого, нужно стандартизировать две вещи: язык общения устройств и физический способ их соединения. Для того, чтобы сообщения можно было сохранять, а впоследствии воспроизводить, нужен также стандартный формат их хранения.

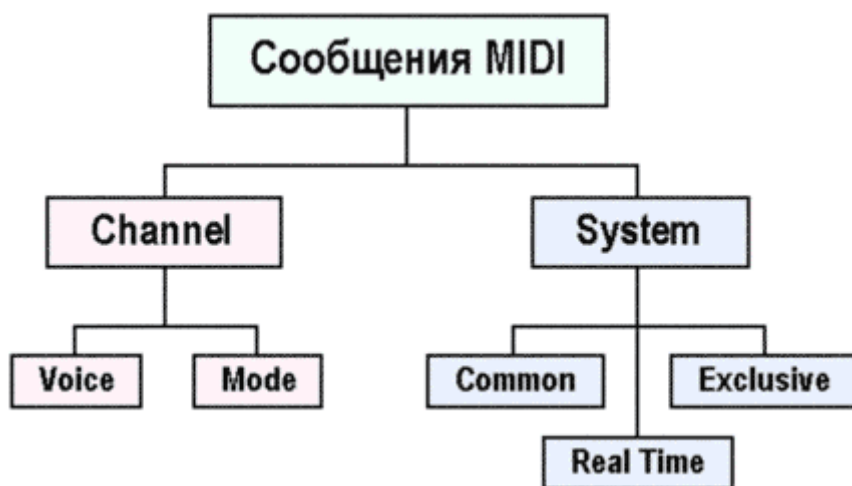
Протокол MIDI состоит из трех частей: спецификации формата данных, аппаратной спецификации интерфейса и спецификации формата хранения данных.

Начнем разговор о MIDI с первой части, то есть с языка, на котором общаются устройства.

Язык MIDI. Язык MIDI состоит только из команд управления и параметров этих команд. Ничего другого по MIDI-кабелю не передается. Любые, даже такие непростые музыкальные моменты, как, например, игра баса легато с "подъездом" к каждой ноте на сильной доле такта, — это всего лишь результат обработки команд тон-генератором. Команды в языке MIDI называются сообщениями.

Сообщения логично разделить на два основных типа: одни управляют звукообразованием, то есть говорят, например, какую ноту и как громко играть, вторые выполняют служебные функции, вроде изменения настроек тон-генератора и синхронизации. Поскольку извлечение звуков происходит в MIDI на определенном канале, сообщения первого типа называются сообщениями канала (Channel Messages). Сообщения второго типа называются системными (System Messages). Сообщения

канала всегда относятся к какому-либо MIDI-каналу и обрабатываются тон-генератором только в том случае, если он настроен на прием данного канала. Выше говорилось, что соединение нескольких тон-генераторов через разъем MIDI Thru, при соответствующих настройках инструментов, позволяет передавать "каждому свое". Системные сообщения предназначены всем устройствам в MIDI-системе, независимо от того, на прием каких каналов они настроены.



Сообщения канала делятся, в свою очередь, на голосовые (Channel Voice Messages) и сообщения режима канала (Channel Mode Messages). Системные сообщения делятся на общесистемные (System Common Messages), сообщения реального времени (System Real Time Messages) и эксклюзивные (System Exclusive Messages).

Голосовые сообщения канала посылают в тон-генератор информацию об управлении звуком. Они "докладывают" тон-генератору о том, что сейчас делает исполнитель — нажимает клавишу, крутит колесо модуляции, двигает фейдер или отпускает педаль. То есть голосовые сообщения описывают действия исполнителя в цифровой форме. Тон-генератор, получая эти сообщения, анализирует их и реагирует на них примерно так: "Ага, это нажатие клавиши До первой октавы — нужно воспроизвести ноту; это поворот звуковысотного колеса — нужно сместить у ноты высоту тона; а это — неизвестное мне сообщение, ничего делать не буду".

Общесистемные сообщения выполняют несколько разнородных задач. В их числе — синхронизация MIDI- и аудиоустройств (например, магнитофонов) посредством протокола MIDI Time Code (MTC), передача позиции песни, выбор песни и даже запрос

на подстройку осцилляторов синтезатора.

Системные сообщения реального времени предназначены для синхронизации MIDI-устройств, например, секвенсоров и драм-машин, по протоколу MIDI Clock. Сюда же входят сообщения начальной инициализации ("сброса") устройства и предотвращения некорректной работы (выключения зависших нот). Сообщения реального времени отличаются от всех других тем, что имеют наивысший приоритет передачи, то есть, например, легко могут вклиниться между частями другого сообщения. Это и понятно — ритмическая точность дороже всего.

Системные эксклюзивные сообщения (обозначаемые для краткости SysEx) — это своеобразная палочка-выручалочка для производителей устройств. Сегодня на рынке существует множество типов оборудования, и каждый конкретный прибор имеет свои специфические возможности. Организации MMA и JMSC могли пойти по пути постоянных обновлений спецификации MIDI: появилась какая-то новая функция, скажем, в дымовой машине — выделить под нее новое сообщение и занести в стандарт. Понятно, что это тупиковый путь. За всем не уследишь, а абсолютной совместимости между устройствами все равно не добиться (да и какая совместимость может быть у дымовой машины и синтезатора?). Поэтому было решено дать возможность каждому производителю определять свои собственные (эксклюзивные, исключительные) сообщения, даже под конкретное устройство.

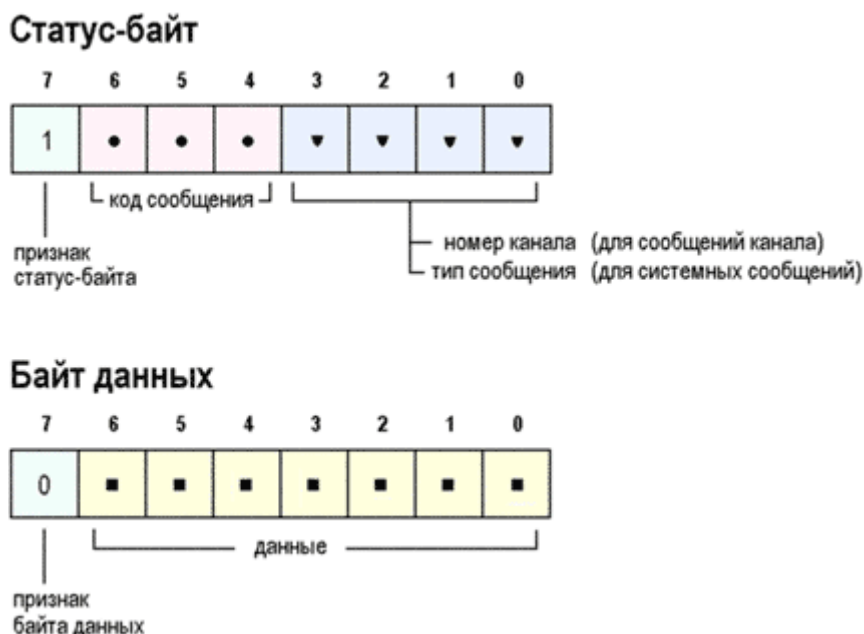
Несколько сообщений SysEx являются универсальными и поддерживаются устройствами разных производителей. Эти сообщения применяются для настройки таких параметров тон-генератора, как общая громкость и панорама, а также для реализации дополнительных протоколов в рамках MIDI, например, протокола передачи семплов (MIDI Sample Dump Standard), управления сценическим светом и пиротехническими устройствами (MIDI Show Control), управления транспортными функциями оборудования (MIDI Machine Control).

Принцип кодирования сообщений. Компоненты сообщений в протоколе MIDI представлены байтами. Компонент, описывающий тип сообщения, называется статус-

байтом, компонент, уточняющий сообщение — байтом данных. Если уточняющей информации много, она может быть представлена несколькими байтами данных. Таким образом, каждое MIDI-сообщение состоит из одного статус-байта и, если необходимо, одного или нескольких байтов данных. Передаются эти сообщения по MIDI-кабелю в том же порядке — сначала статус-байт, затем байты данных. Количество байтов данных жестко закреплено за каждым сообщением. Так что, если, например, тон-генератор получил статус-байт "нажата клавиша", он ждет за ним два байта данных, первый из которых содержит номер нажатой клавиши, второй — скорость нажатия.

Для системных эксклюзивных сообщений сделано исключение — их длина жестко не задается. Она определяется специальным статус-байтом, который помещается в конец сообщения.

В протоколе MIDI каждый статус байт содержит в старшем значащем разряде единицу, а каждый байт данных — ноль. Так что один бит из байта расходуется по служебной необходимости. Для кодирования сообщений и передачи данных остаются только семь бит.



Получается такая картина: старший бит статус-байта занят, так как должен всегда содержать единицу, четыре младших байта — тоже, так как содержат номер MIDI-канала. На кодирование типа сообщения остается всего три бита, которыми можно

закодировать восемь сообщений. Семь из них отвели под сообщения канала, а последнее, все биты которого единичные (111), сделали признаком системного. Системные сообщения не требуют номера канала, поэтому младшие четыре бита статус-байта могут быть использованы для кодирования конкретного типа сообщения (всего их получается 16).

Сообщения канала											
номер	код										
					номер канала						
1	1	0	0	0	n	n	n	n			
2	1	0	0	1	n	n	n	n			
3	1	0	1	0	n	n	n	n			
4	1	0	1	1	n	n	n	n			
5	1	1	0	0	n	n	n	n			
6	1	1	0	1	n	n	n	n			
7	1	1	1	0	n	n	n	n			
Системные сообщения											
номер	код										
1	1	1	1	1	0	0	0	0			
2	1	1	1	1	0	0	0	1			
3	1	1	1	1	0	0	1	0			
4	1	1	1	1	0	0	1	1			
5	1	1	1	1	0	1	0	0			
6	1	1	1	1	0	1	0	1			
7	1	1	1	1	0	1	1	0			
8	1	1	1	1	0	1	1	1			
9	1	1	1	1	1	0	0	0			
10	1	1	1	1	1	0	0	1			
11	1	1	1	1	1	0	1	0			
12	1	1	1	1	1	0	1	1			
13	1	1	1	1	1	1	0	0			
14	1	1	1	1	1	1	0	1			
15	1	1	1	1	1	1	1	0			
16	1	1	1	1	1	1	1	1			

Уровни кодирования

можно выделить два множества сообщений или, иначе, два уровня кодирования. Доступ к первому — прямой, ко второму — только через посредника.

Следующая таблица показывает сообщения первого уровня (рис. 4). Статус-байты даны в двоичном, десятичном и шестнадцатеричном виде. Младшая половина статус-байта (4 бита) содержит номер канала, которому адресовано сообщение. Так, например статус-байты в диапазоне от 10010000 (0x90) до 10011111 (0x9F) означают одно и то же сообщение Note On, в диапазоне от 11000000 (0xC0) до 11001111 (0xCF) — сообщение Program Change и т. д. То есть, статус-байт, соответствующий одному типу сообщения, может быть представлен в шестнадцати вариантах, в зависимости от номера канала. В MIDI, как и во всем компьютерном мире, отсчет ведется с нуля. При передаче сообщений первый канал кодируется как 0000, второй как 0001, а шестнадцатый — как 1111 (десятичное 15, шестнадцатеричное 0xF).

Статус-байт			Сообщение	Байт данных 1	Байт данных 2
Hex	Bin	Dec			
8n	1000nnnn	128..143	Note Off	номер ноты	динамика
9n	1001nnnn	144..159	Note On	номер ноты	динамика
An	1010nnnn	160..175	Polyphonic Key Pressure	номер ноты	давление
Bn	1011nnnn	176..191	Control Change	номер контроллера	значение контроллера
Cn	1100nnnn	192..207	Program Change	номер программы	—
Dn	1101nnnn	208..223	Channel Pressure	давление	—
En	1110nnnn	224..239	Pitch Wheel Change	LSB	MSB

Основной "целевой группой" MIDI являются инструменты клавишного типа (синтезаторы), поэтому набор наиболее часто употребляемых сообщений вполне обоснован. Это, прежде всего, сообщения

- о нажатии и отпускании клавиши (Note On / Note Off),
- Послекасание (сила нажатия на клавишу) (Pressure),
- вращения звуковысотного колеса (Pitch Wheel Change)
- смены программы (Program Change).

Извлечение звуков MIDI является выраженным клавишно-ориентированным протоколом, поэтому процесс извлечения звуков кодируется двумя простыми

сообщениями — взять ноту (Note On) и снять ноту (Note Off). Исполнитель при нажатии клавиши задает сразу три параметра: *момент начала звучания, динамику и высоту тона*. Длительность звука определяется по моменту отпускания клавиши.

В MIDI используется равномерно-темперированный строй и 128 нот различной высоты (с номерами от 0 до 127). Частота нот задается с помощью номера. Нота с номером 60 — всегда До первой октавы (частота 261 Гц):



Полный частотный диапазон, покрываемый таким образом, превышает 10 октав (от 8 Гц до 12543 Гц), что более чем достаточно. Для перевода номера ноты в частоту можно использовать формулу

Зависимость частоты ноты от ее MIDI-номера

$$f_N = 440 \cdot 2^{\frac{N-69}{12}},$$

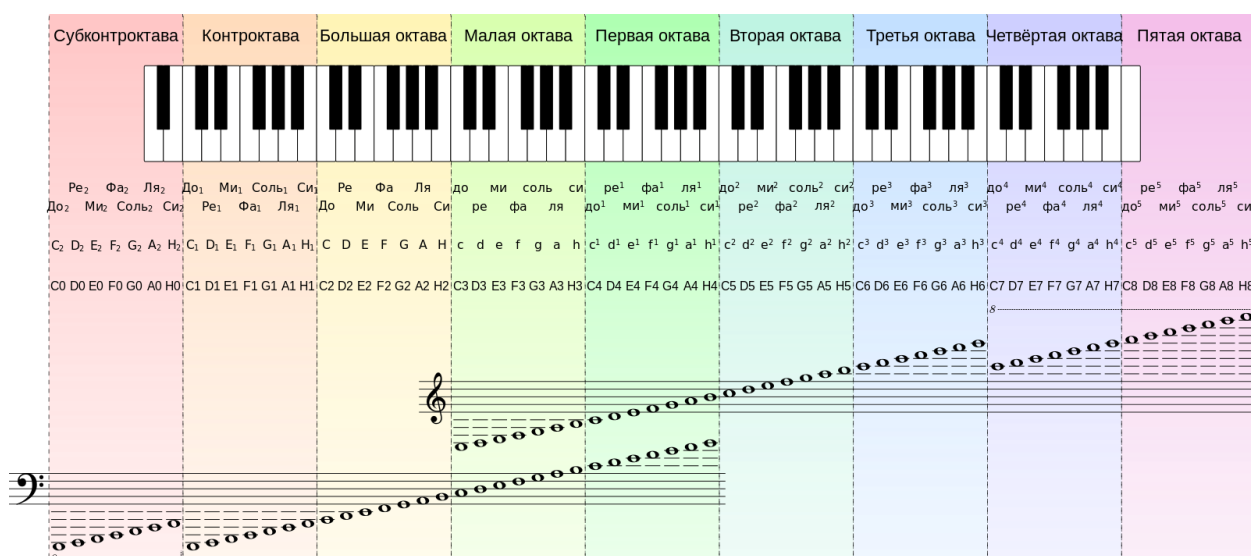
где N — номер ноты;
 f_N — частота ноты в Гц.

$$\sqrt[12]{2} = 1.0594631$$

Частоты в герцах (интервал от До первой октавы в полутонах)											
Октава → Нота ↓	Суб-контр	Контр	Большая	Малая	1	2	3	4	5	6	7
C	16,352 (-48)	32,703 (-36)	65,406 (-24)	130,81 (-12)	261,63 (±0)	523,25 (+12)	1046,5 (+24)	2093,0 (+36)	4186,0 (+48)	8372,0 (+60)	16744,0 (+72)
C# / D $\flat$	17,324 (-47)	34,648 (-35)	69,296 (-23)	138,59 (-11)	277,18 (+1)	554,37 (+13)	1108,7 (+25)	2217,5 (+37)	4434,9 (+49)	8869,8 (+61)	17739,7 (+73)
D	18,354 (-46)	36,708 (-34)	73,416 (-22)	146,83 (-10)	293,66 (+2)	587,33 (+14)	1174,7 (+26)	2349,3 (+38)	4698,6 (+50)	9397,3 (+62)	18794,5 (+74)
D# / E $\flat$	19,445 (-45)	38,891 (-33)	77,782 (-21)	155,56 (-9)	311,13 (+3)	622,25 (+15)	1244,5 (+27)	2489,0 (+39)	4978,0 (+51)	9956,1 (+63)	19912,1 (+75)
E	20,602 (-44)	41,203 (-32)	82,407 (-20)	164,81 (-8)	329,63 (+4)	659,26 (+16)	1318,5 (+28)	2637,0 (+40)	5274,0 (+52)	10548 (+64)	21096,2 (+76)
F	21,827 (-43)	43,654 (-31)	87,307 (-19)	174,61 (-7)	349,23 (+5)	698,46 (+17)	1396,9 (+29)	2793,8 (+41)	5587,7 (+53)	11175 (+65)	22350,6 (+77)
F# / G $\flat$	23,125 (-42)	46,249 (-30)	92,499 (-18)	185,00 (-6)	369,99 (+6)	739,99 (+18)	1480,0 (+30)	2960,0 (+42)	5919,9 (+54)	11840 (+66)	23679,6 (+78)
G	24,500 (-41)	48,999 (-29)	97,999 (-17)	196,00 (-5)	392,00 (+7)	783,99 (+19)	1568,0 (+31)	3136,0 (+43)	6271,9 (+55)	12544 (+67)	25087,7 (+79)
G# / A $\flat$	25,957 (-40)	51,913 (-28)	103,83 (-16)	207,65 (-4)	415,30 (+8)	830,61 (+20)	1661,2 (+32)	3322,4 (+44)	6644,9 (+56)	13290 (+68)	26579,5 (+80)
A	27,500 (-39)	55,000 (-27)	110,00 (-15)	220,00 (-3)	440,00 (+9)	880,00 (+21)	1760,0 (+33)	3520,0 (+45)	7040,0 (+57)	14080 (+69)	28160,0 (+81)
A# / B $\flat$	29,135 (-38)	58,270 (-26)	116,54 (-14)	233,08 (-2)	466,16 (+10)	932,33 (+22)	1864,7 (+34)	3729,3 (+46)	7458,6 (+58)	14917 (+70)	29834,5 (+82)
B	30,868 (-37)	61,735 (-25)	123,47 (-13)	246,94 (-1)	493,88 (+11)	987,77 (+23)	1975,5 (+35)	3951,1 (+47)	7902,1 (+59)	15804 (+71)	31608,5 (+83)

Примечание: Иногда нота Си обозначается как «Н» вместо «В».

С номерами нот работать неудобно, но и музыкальный способ наименования (вроде $d3$ — Ре третьей октавы или $Cis1$ — До диез контроктавы) не подходит; еще и потому, что в разных странах применяются разные обозначения. Спецификация MIDI по этому поводу молчит, а большинство производителей используют такую систему: октава указывается после буквенного обозначения ноты, но первая октава (с музыкальной точки зрения) считается третьей.



Динамика

Динамику извлечения звука называют по-разному: просто динамика, скорость нажатия, громкость. Наиболее корректно в случае клавишного инструмента говорить о *скорости*, поскольку именно скорость нажатия клавиши определяет громкость звучания ноты. Сила нажатия не всегда однозначно соответствует громкости — так, можно довольно сильно давить на клавишу фортепиано, но при этом инструмент будет звучать тихо. Это связано с особенностями механики клавишных.

В спецификации MIDI также используется термин "скорость нажатия", Velocity (от англ. "velocity" — "скорость"). Но для удобства (в ущерб корректности) будем называть "скорость нажатия" одним словом — "динамика".

Разработчики MIDI решили кодировать динамику одним байтом данных, и поэтому в MIDI есть 128 уровней Velocity. Можно задать динамики как целой партии (или громкости канала), так и для отдельно взятой ноты.

Использование значения динамики (за исключением нулевого) спецификацией жестко не определено и зависит от конкретного тон-генератора. Как правило, динамика управляет блоком VCA инструмента — в зависимости от динамики изменяется время или уровень атаки, а, следовательно, и общая громкость звука. Часто динамика управляет параметрами фильтра, делая звук ярче при громкой игре. В семплерах может применяться раскладка по динамике, когда на одну и ту же ноту вешаются несколько семплов, полученных при записи с разными нюансами исполнения. В зависимости от динамики воспроизводится нужный семпл, либо результат микширования двух смежных семплов (для маскирования их разного спектрального состава). Это позволяет довольно успешно имитировать акустические инструменты, звук которых может принципиально отличаться по тембру при разной динамике исполнения.

Возникает вопрос: как должна зависеть громкость сигнала на выходе осциллятора от значения Velocity? Спецификация MIDI рекомендует тон-генераторам использовать логарифмическую зависимость в соответствии с формулой

Зависимость уровня сигнала от Velocity

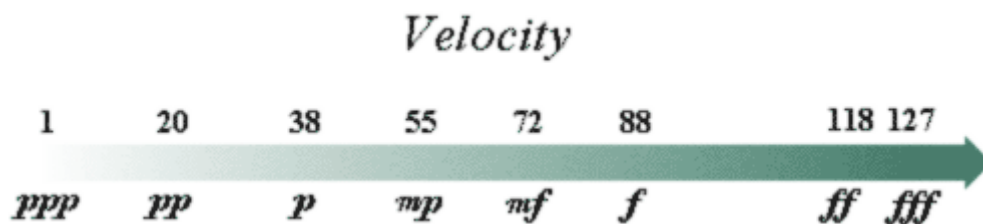
$$L_{dB} = 40 \lg \left(\frac{Velocity}{127} \right)$$

Где L_{dB} — максимальный уровень на выходе осциллятора при заданной огибающей сигнала; обычно точка максимального уровня совпадает с моментом завершения стадии атаки. Зависимость амплитуды сигнала (в процентах) от значения динамики можно найти по формуле:

Зависимость амплитуды сигнала от Velocity

$$A_{\%} = 10^{2 \lg \left(\frac{Velocity}{127} \right)} \cdot 100\%$$

При максимальной динамике ($Velocity = 127$) этот уровень равен 0 дБ (100%), при динамике 1 — -84,15 дБ (0,000062%). Значение динамики 0 используется для служебных целей (см. далее) и в формуле неприменимо. Таким образом, MIDI рекомендует использовать для отдельного звука динамический диапазон в 84 дБ. Многие тон-генераторы не следуют этой рекомендации буквально, ведь для большинства звуков такой огромный диапазон просто не нужен. Но все же зависимость, близкая к логарифмической, обычно выдерживается на каком-либо отрезке динамики. MIDI также рекомендует шкалу соответствия музыкальных динамических оттенков и Velocity, показанную на рис.



Dynamics' Note Velocity

Dynamic	Velocity*	Voice
<i>ppp</i>	16	Whispering
<i>pp</i>	33	Almost at a whisper
<i>p</i>	49	Softer than speaking voice
<i>mp</i>	64	] Speaking voice
<i>mf</i>	80	
<i>f</i>	96	Louder than speaking
<i>ff</i>	112	Speaking loud
<i>fff</i>	127	Yelling



Decrescendo
(diminuendo)



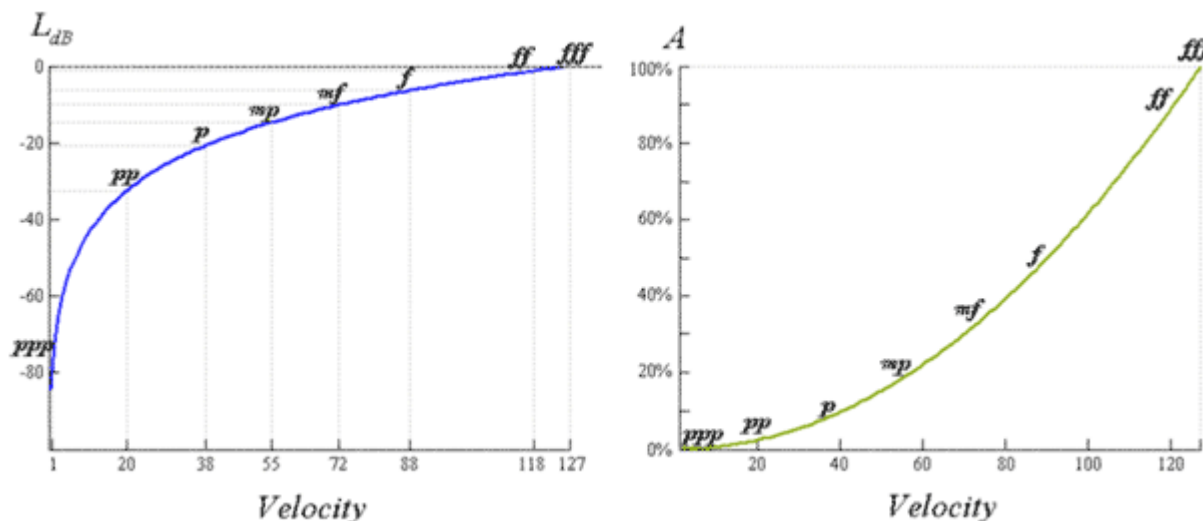
Crescendo



Accent

*Note velocity adopted from Logic Pro

Обращает на себя внимание большой промежуток между forte (*f*, Velocity = 88) и fortissimo (*ff*, Velocity = 118). На следующих графиках (рис. 31) показана рекомендуемая логарифмическая зависимость: уровень в децибелах от динамики (левый график) и относительный уровень (в процентах) от динамики (правый график). Для удобства на графиках отмечены музыкальные динамические оттенки.



Отпускание клавиши

Сообщение **Note Off** дает команду снять ноту заданной высоты. Первый байт данных содержит MIDI-номер ноты, второй — динамику ее снятия. Снятие ноты заставляет тон-генератор перейти к стадии затухания (Release) огибающей. Таким образом, после

снятия пальца с клавиши нота звучит еще ровно столько, сколько длится этот участок. Правда, если отпустить клавишу при нажатой педали Sustain (контроллер CC#64, Hold Pedal, см. далее), то звуковой генератор, хотя и получит сообщение Note Off, но участок Release не включит. Нота продолжит звучание на участке поддержки (Sustain) до тех пор, пока педаль не будет отпущена. При отпуске педали генератор, помня о том, что сообщение Note Off уже приходило, включит участок Release автоматически, и нота скоро прекратит звучание.

Послекасание

Послекасание (Aftertouch) — это давление, прикладываемое к клавише после ее нажатия. Послекасание позволяет имитировать приемы звукоизвлечения, несвойственные клавишному инструменту, например, вибрато или тремоло, причем довольно естественным и удобным способом — нажав клавишу, можно непрерывно изменять давление на нее.

Это давление измеряют сенсоры клавиатуры. Делать сенсор под каждую клавишу — дорогое удовольствие, поэтому большинство клавиатур имеют один сенсор, который измеряет общее давление. Исходя из этого, спецификация определяет два варианта послекасания: монофоническое (или послекасание канала, Channel Pressure) — это сообщение генерируют клавиатуры с одним-единственным сенсором, и полифоническое (Polyphonic Key Pressure) — для клавиатур с сенсором на каждую клавишу. Бывают клавиатуры, которые не генерируют послекасания вообще.

Понятно, что для изменений в звуке генератор должен понимать сообщения послекасания и как-то на них реагировать. Обычно изменяются уровень поддержки (Sustain) на огибающей громкости или фильтра, или глубина/скорость модуляции, производимой генератором низкой частоты. Спецификация не определяет конкретной реакции устройства на сообщения послекасания, и многие синтезаторы их игнорируют.

Послекасание канала менее информативно и не имеет такой гибкости и выразительности, как полифоническое. Так, если вы дожмете одну клавишу больше другой на устройстве, передающем послекасание канала, оба давления будут

усреднены, словно вы дожали обе клавиши с одинаковой силой. Если в аккорде вы хотите выделить басовую ноту, ничего не выйдет — эффект будет применен ко всем нотам, звучащим на канале.

Обычно клавиатура посылает один вид послекасания. Но бывают клавиатуры (с сенсором на каждую клавишу), которые могут посылать оба вида сообщений сразу: полифоническое послекасание формируется как обычно, а послекасание канала вычисляется как среднее от полифонического по каждой клавише.

Послекасание — типичный непрерывный контроллер низкого разрешения, с характерной особенностью: оба типа сообщений довольно сильно загружают канал передачи данных. Давить на клавиши с постоянным усилием получается редко, а при любом изменении давления передается новая порция байт. Послекасание канала загружает канал передачи значительно меньше, чем полифоническое, так как производит меньше байт в единицу времени.

Звуковысотное колесо Звуковысотное колесо (Pitch Wheel) используется для плавного изменения высоты тона звучащей ноты. Процесс изменения высоты называется питч-бендом (Pitch Bend), а соответствующее MIDI-сообщение — Pitch Wheel Change (его часто и называют Pitch Bend). Сообщение об изменении высоты относится сразу ко всему MIDI-каналу, поэтому нельзя, например, сыграть на одном канале аккорд, а сдвинуть высоту тона только у его басовой ноты.

Pitch Wheel Change — сообщение высокого разрешения, то есть позволяет использовать 16384 значений изменения высоты, и относится к контроллерам непрерывного типа. При движении колеса высота ноты скачкообразно изменяется во времени, но благодаря маленьким шагам (равным $1/16384$ диапазона изменения) эти скачки обычно незаметны, и создается ощущение плавного изменения высоты.

Модуляция

Сообщение Modulation (CC#1/CC#33) задает глубину частотной модуляции (вибрато) и управляет соответствующим LFO-блоком тон-генератора. В синтезаторах управление модуляцией происходит обычно с помощью колеса, либо рычага на передней панели.

Диапазон, в котором изменяется глубина модуляции, спецификация оставляет на усмотрение тон-генератора.

Громкость и экспрессия

Итак, параметра Velocity для выстраивания баланса нескольких партий недостаточно. На помощь приходит сообщение MIDI Volume (контроллер CC#7/CC#39), которое задает громкость звучания на MIDI-канале. Его действие можно сравнить с фейдером канала в микшерском пульте. С какой бы динамикой не извлекал звуки исполнитель, уровень партии в миксе будет зависеть от положения фейдера. Говоря языком MIDI — какая бы скорость нажатия не использовалась, выходная громкость партии будет зависеть от значения MIDI Volume.

В современных тон-генераторах тембр при увеличении динамики, как правило, изменяется (становится ярче). Применяя совместно Velocity и MIDI Volume, можно получить любые сочетания "яркость/громкость". Например, максимальное значение Velocity при небольшой громкости канала создает впечатление очень яркого по тембру, но тихого инструмента.

В мультитембральном тон-генераторе после отстройки относительного баланса партий (с помощью сообщения MIDI Volume) может потребоваться регулировка общей громкости устройства (уровня мастер-выхода). Для этого чаще всего используется сообщение SysEx Master Volume (поговорим в следующей статье), иногда — значение громкости одного из каналов или контроллера типа General Purpose.

Большинство устройств игнорируют LSB в сообщении MIDI Volume и позволяют производить только грубую настройку (128 уровней), посредством MSB (контроллер CC#7). Некоторые старые синтезаторы на сообщение MIDI Volume не реагируют.

Бок о бок с сообщением громкости канала идет сообщение экспрессии (Expression, CC#11 / CC#43). Его название может сбить с толку, поскольку прямой перевод с английского (expression — выразительность) — довольно расплывчатое понятие. Многие в этой связи склонны ожидать от контроллера экспрессии какую-либо тонкую

настройку пэтча (например, посредством фильтров) или других чудес. На самом деле все гораздо проще.

Слово Expression как музыкальный термин соответствует понятию "нюанс" русской музыкальной терминологии, которое означает, прежде всего, динамический нюанс исполнения — форте, пиано и т. д. Таким образом, контроллер экспрессии предназначен для дополнительного управления громкостью MIDI-канала. Как и в случае с сообщением MIDI Volume, большинство тон-генераторов допускают только грубую настройку экспрессии (128 уровней) посредством MSB (контроллер CC#11).

При максимальном значении экспрессии громкость канала равна текущему значению MIDI Volume. При уменьшении экспрессии начинается уменьшение результирующей громкости. Когда экспрессия равна нулю, громкость канала также равна нулю; когда экспрессии "половина", то и громкость канала равна половине от текущего значения MIDI Volume.

Гораздо проще управлять динамикой внутри партии с помощью сообщения экспрессии, а с помощью MIDI Volume балансировать партию со всеми остальными. Тогда в нашем случае задача упрощается: нужно только увеличить общую громкость канала, а кривая нарастания останется неизменной.

Исходя из этого, протокол MIDI рекомендует использовать сообщение MIDI Volume только для общей балансировки партий, один раз в самом начале аранжировки, а динамические изменения в процессе звучания производить с помощью сообщения экспрессии.

Панорама

Сообщение Pan (контроллеры CC#10/CC#42) устанавливает панораму на MIDI-канале, то есть определяет положение партии в стереобазе. Сообщение имеет смысл только для тех тон-генераторов, которые имеют стереовыход или несколько выходов, работающих как стереопары.

В зависимости от конкретной архитектуры тон-генератора это сообщение может быть

регулятором панорамы или баланса. Например, если структура пэтча такова, что в блок VCA инструмента приходит моносигнал, то под воздействием сообщения Pan синтезатор просто делит моносигнал на левый и правый канал в нужной пропорции, и с выхода блока снимается псевдостереосигнал. Но во многих тон-генераторах используются стереосемплы (или задействованы два осциллятора, или несколько слоев с установленным положением слоя в стереобазе на уровне пэтча). Так как в этом случае сигнал приходит на вход VCA уже стереофоническим, то сообщение будет работать как регулятор баланса, то есть задавать относительный уровень левого и правого каналов.

Зависимость уровней левого и правого каналов от значения панорамы

$$L_{dBLeft} = 20 \lg \sqrt{1 - \frac{P}{127}}$$

$$L_{dBRight} = 20 \lg \sqrt{\frac{P}{127}},$$

где P — значение Pan (CC#10).

Синтез

В каждом синтезаторе применяется своя схема синтеза, своя конструкция компонентов (осцилляторов, фильтров, усилителей и т. д.), свои варианты их использования. Причем в продукции одной и той же фирмы эти варианты часто отличаются. Поэтому спецификация определяет только самые общие, наиболее характерные для большинства тон-генераторов параметры. Например, управление длительностью фаз в генераторе огибающей. Если предложенные сообщения производителя не удовлетворяют, он всегда может использовать SysEx или NRPN.

Стандарты MIDI

В начале существования интерфейса MIDI у каждого изготовителя были собственные методики организации звуков и каналов. Некоторые устройства использовали один банк данных из 128 сэмплов, другие делили их на кусочки поменьше. Звуки пианино и клавишных могли храниться рядом или быть случайным образом разбросаны по всей памяти устройства. Это вызывало постоянную путаницу, а композиторам, сочинявшим музыку для игр или других приложений, нужно было писать варианты для каждой системы.

General MIDI

Спецификация General MIDI (сокращенно GM) упорядочила этот хаос, создав систему нумерации сэмплов. Теперь, независимо от того, кто сделал ваш синтезатор или звуковую карту, сэмпл номер 1 — это акустический рояль, а номер 128 — звук выстрела. Благодаря GM вам не надо беспокоиться о том, что ваше творение будет (или не будет) воспроизводиться правильно. Почти любые существующие синтезаторы, звуковые карты или другие устройства воспроизведения сэмплов (и аппаратные, и программные) поддерживают GM, а многие и расширяют его. Это же касается операционных систем Mac OS и Windows.

GM использует каналы 1—9 и 11—16 для таких инструментов, как клавишные, гитары и духовые (канал 10 закреплен за ударными). Для этих инструментов настройка стандартизирована таким образом, чтобы воспроизведение клавиши под номером 60 звучало как до первой октавы.

Звуки размещены в более или менее логичном порядке: например, звуки пианино сгруппированы в банках с 1 по 9, за ними следуют звуки хроматических ударных, такие как маримба, в каналах с 9 по 16. Звуки басовых можно найти в каналах с 33 по 40. Хотя спецификация GM определяет, какой звук будет вызываться, она ничего не говорит о том, как он должен воспроизводиться.

Соответственно качество и характер отдельных звуков различаются у разных производителей или даже у различных устройств одного изготовителя. Некоторые звуки GM, например акустического пианино, известны своей сложностью качественного исполнения. Другие же, например звук блок-флейты (№ 75), очень просто воссоздать даже на самой дешевой звуковой карте

Канал 10 зарезервирован за звуками ударных, таких как наборы барабанов или ручные ударные инструменты. В отличие от хроматических звуков других каналов, каждая нота канала № 10 используется для выбора конкретного ударного инструмента. Это позволяет играть партию ударных прямо на клавишном пульте. Например, используйте ноты № 60 и 61 (среднее до и ре на октаву выше) для ритма с высоким и низким бонго. Некоторые GM-пульты и модули расширяют идею единого набора барабанов. Используя сообщения о выборе банка данных, вы сможете отдавать предпочтение тому или иному комплекту, разработанному для джаза, тяжелого рока, электронной музыки и другого.

General MIDI — это нечто больше, чем просто список звуковых банков. Спецификации требуют, чтобы совместимые звуковые модули (и аппаратные, и программные) распознавали все 16 MIDI-каналов и использовали любой из 24 динамически распределяемых голосов для одновременного воспроизведения звуков ударных и мелодии, иногда разделяемых на 8 голосов для ударных и 16 — для звуков мелодии. GM также определяет, как модули будут отвечать на конкретные сообщения контроллера, и некоторые другие полезные мелочи.

Главное, что можно создать песню на своей системе и знать, что она будет правильно воспроизводиться на любой GM. Так как файлы MIDI очень маленькие по сравнению с файлами цифрового аудио, MIDI идеально подходит для использования в Сети.

Roland GS

GS начал развиваться в 1991 году сразу за GM и неоднократно расширялся в связи с выпуском новых продвинутых моделей. XG от Yamaha активно продвигался с 1996 года, и при его разработке учитывался опыт фирмы Roland.

Стандарт GS имеет обратную совместимость с General MIDI, то есть любая MIDI-композиция, созданная для GM-устройств, может быть успешно воспроизведена и на GS-устройствах. Все GS должны соответствовать требованиям к GM-устройствам. Кроме того, для них предусмотрены некоторые дополнительные условия:

минимальный набор из 226 инструментов (вместо 128 в GM), причем для выбора инструментов, не входящих в базовый набор, используется переключение банков с помощью контроллера № 0;

несколько (до 128) различных 61-звучных наборов ударных;

поддержка как минимум 20 контроллеров;
возможность динамической редакции по MIDI восьми звуковых параметров (время атаки и затухания, частота среза фильтра, скорость вибрато и пр.);
возможность динамической редакции по MIDI пяти параметров для ударных тембров (высота, пространственная локализация, а также уровень громкости, реверберации и хоруса);
широкие возможности редактирования звучания с помощью исключительных системных MIDI-сообщений.

Стандарт GS не получил такого широкого распространения, как General MIDI, однако некоторые музыканты все же отдают ему предпочтение. Использовать у себя Roland GS вы можете в том случае, если у вас есть соответствующий синтезатор или звуковой модуль этой фирмы

Yamaha XG

А в сентябре 1994 года появился формат XG (Extended General MIDI), разработанный фирмой Yamaha, чтобы расширить и усложнить стандарт GM (люди вообще любят все усложнять). XG полностью обратно совместим с GM и в то же время обеспечивает большой набор дополнительных элементов: новые инструменты и редактирование их, несколько сотен эффектов и множество других возможностей, которые используются мультимедийными приложениями. Вот почему многие компании — производители музыкальной аппаратуры приняли данный стандарт и выпускают технику, которая позволяет его применять.

Прежде всего спецификация XG поддерживает гораздо большее количество инструментов, чем GM. Минимальный набор, определенный в спецификации, должен содержать 480 инструментов. Вместе со всеми эффектами цифра получается еще более внушительной — 676 инструментов.

Спецификация XG предусматривает и большие возможности редакции исполнительских параметров по MIDI. Предусмотрено также много эффектов — реверберация (эффект «эха», когда звуковой сигнал смешивается со своими копиями, задержанными на различные временные интервалы), хорус (эффект исполнения одного звука несколькими инструментами) и другие эффекты, коих несколько десятков. Минимальное количество типов реверберации и хоруса — 8, а остальных — 35.

Последние включают в себя как различные варианты и комбинации реверберации и хоруса, так и дополнительные эффекты: вращение, тремоло, «вау-вау», искажения, шумы и т.п. Любой эффект может быть использован как на отдельном MIDI- канале, так и на общем выходе.

Инструменты MIDI

Синтезатор или **Секвэнсор, секвэнсер** (англ. sequencer, от англ. sequence — «последовательность») — устройство (аппаратное или программное) для записи, редактирования и воспроизведения последовательности MIDI-данных, главным образом, формульных ритмических фигур и мелодических фраз (так называемых «паттернов»). В принципе, секвенсор способен записывать любые MIDI-сообщения, например, панорамирование, переключение определённых заранее тембров (патчей), динамические нюансы и другие параметрические характеристики звучания.

Сэмплер - разновидность синтезатора, отличаются тем, что основным фундаментом, на котором строится выходной звук, является короткая цифровая запись (сэмпл).

Конечно, не всегда получается записать сотни сэмплов, которые могут потребоваться, поэтому существуют библиотеки, охватывающие все: от оркестровых инструментов до коротких, ритмичных ударных и шаблонов басовых.

Классификации виртуальных инструментов

Для начала давайте с вами попробуем немного упорядочить виртуальные музыкальные инструменты, то есть классифицировать, взяв за основу тот или иной признак.

По алгоритму синтеза звука. Синтезаторы создают музыкальный тембр на основе вложенных в них алгоритмов, а сэмплеры — тембр на основе загружаемых в них шаблонов — сэмплов, и каждый новый сэмпл расширяет тембровые возможности инструмента. Все виртуальные синтезаторы цифровые по своей сути, но их внутренняя архитектура может опираться на различные идеи. К примеру, программный аналоговый синтезатор с помощью математических моделей имитирует электронные схемы, существующие в реальном прототипе.

По способу запуска можно выделить две категории. Программы, запускаемые сами по себе (standalone): для работы с ними не требуется дополнительных программ или утилит. Вторая категория — это подключаемые модули (плагины, **plug-in**). Они могут появиться только в окне другой программы, называемой хостом. Двумя словами смысл плагинов можно выразить так: на компьютере устанавливается мощная хост-программа

и программный модуль- плагин. При запуске хоста она находит этот плагин и подключает его к себе как компонент. Таким образом программа-хост дополняется новой функцией, которая первоначально в ней не была заложена. Плагины бывают следующих типов: VSTi (PC и Macintosh), DirectX (Windows), DSSI/LADSPA (Linux), MESS (Linux), Audio Units (Macintosh), RTAS (Pro Tools), TDM (Time-division multiplexing) и MAS (MOTU).

По привязке к аппаратной части. Отдельные виртуальные синтезаторы могут работать только в комплексе с определенной звуковой картой — например, компании Digidesign. Плюсом здесь является то, что в этой ситуации все операции (запись и воспроизведение аудио и MIDI, работа программного синтезатора, обработка и микширование аудиопотока и т.д.) проводятся силами специального процессора на звуковой карте и не трогают центральный процессор компьютера. Минус очевиден — на чужом «железе» программа работать не будет. Но большинство виртуальных синтезаторов, к счастью, передоверяют заботу о совместимости с аудиоинтерфейсом операционной системе и соответствующим драйверам.

MIDI-файлы

Одной из трех составляющих протокола MIDI является спецификация формата хранения данных (напомню, что две другие составляющие — это формат сообщений и аппаратная спецификация интерфейса). Он был предложен организацией MMA в конце 1987 года и получил название "Стандартные MIDI-файлы" (Standard MIDI Files, SMF).

Цель MIDI-файлов — обеспечить обмен событиями (то есть MIDI-сообщениями, имеющими метку времени) между различными устройствами и программами. До появления стандартных MIDI-файлов аранжировку, подготовленную в одном секвенсоре, невозможно было загрузить в другой из-за несовместимости форматов.

Поскольку хранение информации непосредственно связано с устройством секвенсоров, остановимся на этом деле подробнее, но лишь в той мере, какая необходима для понимания SMF.

События

Сообщение MIDI, снабженное меткой времени, называется *событием*. Для указания времени могут использоваться разные единицы — тики, внутренние импульсы, время в формате SMPTE и т. п. Важно понять принципиальное отличие события от сообщения. Сообщение "живет" долю миллисекунды реального времени — от момента, когда генерируется источником, до момента, когда поступает к исполнению в приемник. Его можно отловить в процессе передачи по MIDI-кабелю в виде набора импульсов. Событие — это несколько байт информации, записанной в память устройства, на основании которой в будущем, в назначенное время, будет сгенерировано сообщение.

Например, в памяти секвенсора может храниться событие "взять ноту" с меткой времени 100 мс от начала воспроизведения. Вы можете редактировать это событие в двух измерениях: во-первых, изменять параметры самого MIDI-сообщения (в данном случае, высоту или динамику ноты), а во-вторых — перемещать ноту по дорожке, то есть изменять время исполнения сообщения.

События появляются в памяти секвенсора в процессе записи MIDI-сообщений. При

нажатии кнопки Record секвенсор включает аппаратный генератор импульсов (тиков) и начинает "слушать" заданный MIDI-вход. Например, при нажатии клавиши на вход поступает сообщение "взять ноту". Секвенсор смотрит — ага, сообщение пришло на 20-й тик, и записывает его в память с меткой 20. Через пару секунд отпустили клавишу — пришло сообщение "снять ноту", внутренний генератор в этот момент радостно отмахал секвенсору 64 тика. Секвенсор сохраняет сообщение с меткой 64. Вот теперь мы имеем дело с двумя событиями — Note On и Note Off. При воспроизведении снова включается генератор импульсов. Когда подходит 20-й тик, на MIDI-выход секвенсора отправляется сообщение Note On, 64-й тик — Note Off. Мы записали, а затем воспроизвели действия исполнителя!

Очевидно, то же самое можно сделать и в "оффлайне", то есть без необходимости живого исполнения. Щелкнув мышью в нужном месте дорожки (и заранее выбрав длительность ноты), мы построим в памяти секвенсора точь-в-точь предыдущую картину.

Измерение времени

На самом деле, внутри секвенсора никаких "тиков" нет. Есть аппаратный таймер, который генерирует импульсы со строго постоянной частотой (например, каждую микросекунду). Заставлять музыкантов измерять время в микросекундах было бы чудовищным издевательством, как, впрочем, и в других единицах реального времени (секундах, минутах). Музыканты привыкли мыслить в тактах и долях, а время выражать в относительных единицах (длительностях нот), зависящих от текущего темпа.

Только минимальной единицей измерения будет не какая-нибудь 32-я длительность, а условно взятый тик (единица еще мельче, так, что, например, в одной тридцать второй может содержаться 48 тиков). Поскольку еще с классической эпохи темп измеряется количеством четвертей в минуту (BPM, beats per minute), то за основную длительность решили взять четверть и указывать количество тиков, приходящихся на четверть — PPQN (Pulse Per Quarter Note). Чем больше PPQN, тем лучше разрешение секвенсора и тем точнее по времени он может фиксировать сообщения при записи и выдавать их на MIDI-выход при воспроизведении.

Тик — темпозависимая единица: чем быстрее темп, тем короче интервал между тиками в единицах реального времени. Этот интервал можно найти по формуле

$$\Delta t_{[\text{мс}]} = \frac{60\,000}{BPM \cdot PPQN}$$

Большинство секвенсоров позволяют задавать PPQN произвольно — например, от 32 до 1536 тиков на четверть (современные секвенсоры — до 15360 PPQN). Это позволяет осуществить синхронизацию MIDI- и аудиоданных с точностью до семпла. Но как только мы соединяем секвенсор и тон-генератор MIDI-кабелем посредством MIDI-интерфейса, высокое разрешение утрачивает актуальность.

Почему? Напомню, что один байт передается по MIDI-интерфейсу в течение 320 микросекунд. А это значит, что, например, сообщение о взятии ноты (состоящее из трех байт) будет передаваться 960 мкс или почти целую миллисекунду. Теперь представим, что в секвенсоре при темпе 120 BPM и разрешении 2048 PPQN запрограммированы две ноты с интервалом два тика между собой. В единицах реального времени это составляет 488 микросекунд. Так вот: тон-генератор не сможет получить вторую ноту через 488 микросекунд после первой, а реально — только через 960 мкс. Так что исполнит он ее спустя не два тика, а почти четыре.

Отсюда вывод: при работе через MIDI-интерфейс (когда секвенсор и тон-генератор разнесены) разрешение секвенсора более одного тика на 960 микросекунд не имеет смысла. Чтобы узнать, сколько это будет в PPQN, можно воспользоваться формулой на рис. 2.

$$PPQN_{\text{max}} = \frac{60\,000\,000}{BPM \cdot 960}$$

В таблице на рис. 3 для разного темпа показаны значения PPQN, превышать которые бессмысленно.

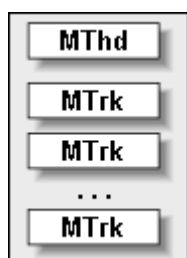
BPM	PPQN_{max}
40	1563
60	1042
80	782
100	625
120	521
140	447
180	348
220	285
240	261

Структура SMF

Стандартные MIDI-файлы состоят из блоков (Chunks). Всего есть два типа блоков: блок заголовка (Header Chunk) и блок трека (Track Chunk). В файле SMF может быть только один блок заголовка и один или более блоков трека. Блок имеет структуру: первые четыре байта — идентификатор, следующие четыре байта — длина блока в байтах, исключая восемь байт тип/длина. Идентификатор блока заголовка — это четыре символа "MThd", блока трека — четыре символа "MTrk". Такая структура позволит в будущем определять новые типы блоков, а незнакомый блок может быть легко проигнорирован на основе его длины. Спецификация SMF предупреждает: "Программы должны быть готовы, встретив блоки неизвестных им типов, игнорировать их".

Блок заголовка содержит общую информацию о файле, блок трека — поток MIDI-сообщений с метками времени. Кроме того, в MIDI-файле сохраняется дополнительная информация, необходимая секвенсорам: темп, размер, тональность, настройки метронома и тому подобное. Эта информация по MIDI-интерфейсу не передается, а образующие ее события называются мета-событиями.

MIDI-файл всегда начинается с блока заголовка, за которым следуют один или более блоков трека:



То есть, любой стандартный MIDI-файл начинается с четырех букв "M", "T", "h", "d". В 1991 году Microsoft предложила "заворачивать" стандартные MIDI-файлы в блок с именем "data" внутри файла RIFF. В результате родился гибрид под названием RMID (расширение RMI). По сути, это все тот же стандартный MIDI-файл, только с приписанной к началу файла шапкой из двадцати байт (рис. 10).

RIFF ID	размер RMID	тип файла	ID блока	размер SMF	SMF
'RIFF'	nn nn nn nn	'RMID'	'data'	nn nn nn nn	'MThd'

20 байт

Лекция 10. Синтез речи

Синтез речи— в широком смысле— восстановление формы речевого сигнала по его параметрам; в узком смысле— формирование речевого сигнала по печатному тексту. Часть искусственного интеллекта.

Синтезом речи— прежде всего называется всё, что связано с искусственным производством человеческой речи.

Синтезатор речи— структура, способная переводить текст/образы в речь, в программном обеспечении или аппаратных средствах.

Голосовой движок— непосредственно система/ядро преобразования текста/команд в речь, это также может существовать независимо от компьютера.

Все способы синтеза речи можно подразделить на группы

- параметрический синтез;
- компиляционный (компилятивный) синтез;
- синтез по правилам;

Параметрический синтез.

Параметрический синтез речи является конечной операцией в вокодерных системах, где речевой сигнал представляется набором небольшого числа непрерывно изменяющихся параметров. Достоинством такого способа является возможность записать речь для любого языка и любого диктора. Качество параметрического синтеза может быть очень высоким (в зависимости от степени сжатия информации в параметрическом представлении).

Вокодер (*voice coder* — кодировщик голоса) — устройство синтеза речи на основе произвольного сигнала с богатым спектром. Изначально вокодеры были разработаны в целях экономии частотных ресурсов радиолинии системы связи при передаче речевых сообщений.

Вокодер представляет собой устройство (или алгоритм), осуществляющее

параметрическое компандирование речевого сигнала.

Компандирование - способ преобразования речевого сигнала, при котором на передающем конце тракта происходит сжатие по одному или нескольким измерениям (частотный диапазон, динамический диапазон, временной интервал), а на приёмном - восстановление первоначального объёма сигнала путём соответствующего расширения. Компандирование включает преобразования: компрессию (сжатие) и декомпрессию (восстановление) речевого сигнала.

Работа вокодера (voice coder) основана на анализе характерных особенностей человеческой речи. На рис. 2 показаны условно частотные характеристики речи как функция от времени.

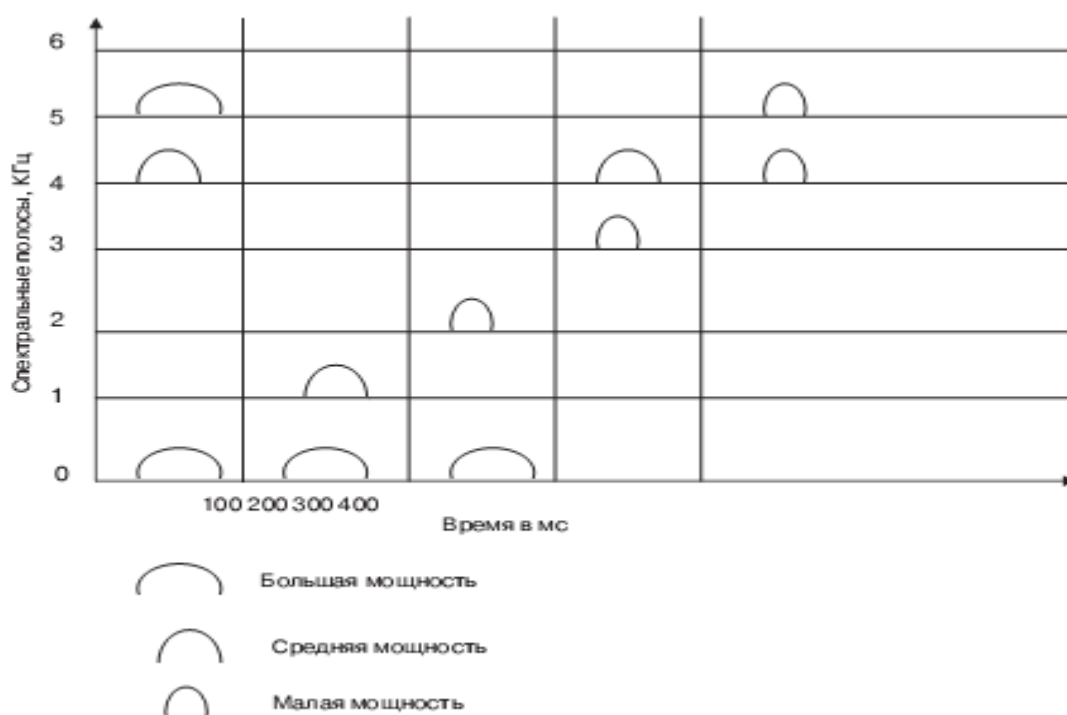


Рис. 2. Пример распределения энергии в частотных диапазонах

На рисунке изображены частотные полосы (от 0 до 1 КГц, от 1 КГц до 2 КГц и т. д.) и распределение энергии по ним при произнесении фразы.

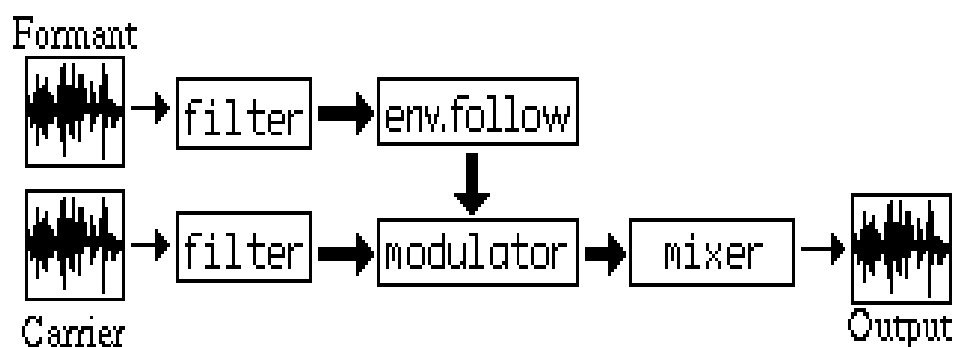
Изображенная "спектрограмма" представляет собой распределение энергии речи в виде функции времени и частоты. Горизонтальная ось представляет время, вертикальная — частоту, уровень энергии условно показан частью синусоиды. Периоды между сменами формант составляют от 10 до 30 мс. Изучение образцов речи показало, что в русском языке содержится 42 фонемы: это 6 гласных звуков и остальные согласные. Чтобы закодировать их номера, достаточно 6 битов.

Как видно из рисунка, энергия распределяется во времени только в некоторых частотных диапазонах и различается по величине. Отдельные пики энергии, возникающие в одном частотном диапазоне, называются фонемами.

Эта картина может изменяться в больших диапазонах, в зависимости от тембра голоса и особенностей произношения, но нам сейчас важно рассмотреть общие закономерности построения. На рисунке видно, что буквы отличаются не только частотным диапазоном, но и структурой. Для каждого звука характерны пики (резонансы) энергии в определенных частотных диапазонах и провалы в других. Частоты, на которых в данный момент возникают комбинации пиков (фонем), называются "частотами формант" или просто "формантами". Гласные и звонкие согласные звуки речи содержат обычно от трех до четырех формант.

Экономия достигается за счёт того, что вместо собственно речевого сигнала передают только значения его определённых параметров, которые на приёмной стороне управляют синтезатором речи. Основу синтезатора речи составляют три элемента:

- генератор тонального сигнала для формирования гласных звуков;
- генератор шума для формирования согласных;
- и система формантных фильтров для воссоздания индивидуальных особенностей голоса.



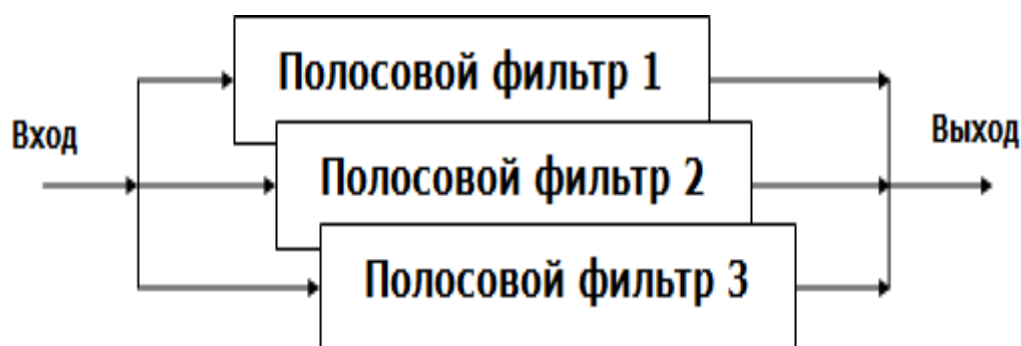
Формантный фильтр является одним из ключевых компонентов в системах синтеза речи и речеподобных сигналов.

Формантный фильтр— система резонансных фильтров, предназначенная для генерации речевого сигнала с заданной фонетической структурой.

В основу структуры формантного фильтра заложена упрощённая модель голосового тракта. В соответствии с моделью, голосовой тракт представляет собой резонатор с несколькими пиками АЧХ, частоты которых определяют вид произносимой фонемы. Эти пики АЧХ получили название форманты. Пример спектра фонемы «А»:



Формантный фильтр создаёт формантные области в спектре входного сигнала с помощью нескольких параллельно соединённых полосовых фильтров. Количество звеньев в схеме определяет порядок формантного фильтра. Схема формантного фильтра третьего порядка:



Чтобы синтезировать речевой сигнал, соответствующий определённой фонеме, необходимо настроить центральную частоту каждого полосового фильтра системы на соответствующую частоту форманты. Таблица частот формант для некоторых фонем (приведённые значения для других голосов, как правило, отличаются):

Фонема	Первая форманта, Гц	Вторая форманта, Гц	Третья форманта, Гц
«и»	270	2300	3000
«е»	400	2000	2550
«а»	660	1700	2400
«у»	640	1200	2400

Входным воздействием для формантного фильтра могут служить различные сигналы с различной окраской тембра. Основными параметрами входного сигнала являются частота повторения и ширина спектра. Частота повторения определяет высоту тона синтезируемой фонемы и лежит в пределах от 200 Гц до 2000 Гц. При этом нижние частоты данного диапазона соответствуют тембру мужского голоса, верхние — женского. Занимаемый входным сигналом диапазон частот должен быть как можно шире. В качестве такого сигнала часто используется импульсная последовательность.

После всех преобразований голос человека становится похожим на "голос робота", что вполне терпимо для средств связи и интересно для музыкальной сферы. Так было лишь в самых примитивных вокодерах 1-й половины XX ст. Современные связные вокодеры обеспечивают высочайшее качество голоса при существенно более сильной степени сжатия в сравнении с упомянутыми выше.

Вокодер как необычный эффект был взят на вооружение электронными музыкантами и впоследствии стал полноценным эффектом благодаря фирмам-изготовителям музыкального оборудования, которые придали ему форму и удобство музыкального эффекта. Вокодер как музыкальный эффект позволяет перенести свойства одного (модулирующего) сигнала на другой сигнал, который называют носителем. В качестве сигнала-модулятора используется голос человека, а в качестве носителя — сигнал, формируемый музыкальным синтезатором или другим музыкальным инструментом. Так достигается эффект «говорящего» или «поющего» музыкального инструмента. Помимо голоса, модулирующий сигнал может быть и гитарой, клавишными, барабанами и вообще любым звуком синтетического и «живого» происхождения. Также нет ограничений и на несущий сигнал. Экспериментируя с моделирующим и несущим сигналом, можно получать совершенно разные эффекты — говорящая гитара, барабаны со звуком фортепиано, гитара, звучащая как ксилофон.

Современные вокодеры можно поделить на:

1. Аппаратные (с которых всё и началось)
2. Виртуальные (появились гораздо позже с развитием компьютерных технологий создания музыки).

В практике компьютерного музыканта значительно чаще используются вокодеры виртуальные, реализованные в виде VST-плагинов (VST-plugins),

Virtual Studio Technology (VST) — формат ресурсозависимых (*native*) плагинов реального времени, которые подключаются к звуковым редакторам и музыкальным редакторам, секвенсорам и т. д.

Применяются подобные решения вокодеров как самостоятельные программы, так и совместно с программой-хостом. В качестве таковой может быть использована любая виртуальная студия, поддерживающая технологию VST, например, **Sound Forge**, **Steinberg Cubase** или **FL Studio**. Программа-хост позволяет подключать собственно сам вокодер и выбирать, откуда будет поступать несущий и модулирующий сигнал (некоторые вокодеры имеют встроенный синтезатор несущего сигнала) — с синтезаторов и семплеров (которые, кстати, имеют тоже формат VST) или с микрофонов и других подключённых к звуковой карте инструментов. А управление несущим сигналом осуществляется посредством MIDI-команд, поступающих из MIDI-секвенсора или MIDI-клавиатуры в VST-плагин (синтезатор или семплер).

Примером виртуальных вокодеров могут служить VST-плагины, такие как: **Steinberg Vocoder**, **Fruity Vocoder**, **Akai DC Vocoder**, **Voctopus**, **AC vocoder**, **Formulator**, **Lpc-vocoder**, **Darkoder**, **Cylonix** (как работающий самостоятельно (*standalone*)). Любой знакомый с его принципиальным устройством может собрать собственный вокодер в любой модульной программе типа **NI Reactor/Generator**, **Max MSP**, **Buzz Composer**.

Компиляционный синтез

Компиляционный синтез сводится к составлению сообщения из предварительно записанного словаря исходных элементов синтеза. Размер элементов синтеза не меньше слова. Очевидно, что содержание синтезируемых сообщений фиксируется объёмом словаря. Основная проблема в компилятивном синтезе — объёмы памяти для хранения словаря. В связи с этим используются разнообразные методы сжатия/кодирования речевого сигнала. Компилятивный синтез имеет широкое практическое применение.

Системами речевого ответа оснащаются разнообразные устройства (от военных самолётов до бытовых устройств) сейчас они находят всё большее применение в повседневной жизни, например, в справочных службах операторов сотовой связи при получении информации о состоянии счета абонента.

Полный синтез речи по правилам

Полный синтез речи по правилам (или синтез по печатному тексту) обеспечивает управление всеми параметрами речевого сигнала и, таким образом, может генерировать речь по заранее неизвестному тексту. В этом случае параметры, полученные при анализе речевого сигнала, сохраняются в памяти так же, как и правила соединения звуков в слова и фразы. Синтез реализуется путём моделирования речевого тракта, применения аналоговой или цифровой техники. Причём в процессе синтезирования значения параметров и правила соединения фонем вводят последовательно через определённый временной интервал, например 5—10 мс. Метод синтеза речи по печатному тексту (синтез по правилам) базируется на запрограммированном знании акустических и лингвистических ограничений и не использует непосредственно элементы человеческой речи. В системах, основанных на этом способе синтеза, выделяется два подхода. Первый подход направлен на построение модели речепроизводящей системы человека, он известен под названием *артикуляторного синтеза*. Второй подход — *формантный синтез по правилам*. Разборчивость и

натуральность таких синтезаторов может быть доведена до величин, сравнимых с характеристиками естественной речи.

Синтез речи по правилам с использованием предварительно запомненных отрезков естественного языка — это разновидность синтеза речи по правилам, которая получила распространение в связи с появлением возможностей манипулирования речевым сигналом в оцифрованной форме. В зависимости от размера исходных элементов синтеза выделяются следующие виды синтеза:

- микросегментный (микроволновый);
- аллофонический;
- дифонный;
- полуслоговой;
- слоговой;
- синтез из единиц произвольного размера.

Обычно в качестве таких элементов используются полуслоги — сегменты, содержащие половину согласного и половину примыкающего к нему гласного. При этом можно синтезировать речь по заранее не заданному тексту, но трудно управлять интонационными характеристиками. Качество такого синтеза не соответствует качеству естественной речи, поскольку на границах сшивки дифонов часто возникают искажения. Компиляция речи из заранее записанных словоформ также не решает проблемы высококачественного синтеза произвольных сообщений, поскольку акустические и просодические (длительность и интонация) характеристики слов изменяются в зависимости от типа фразы и места слова во фразе. Это положение не меняется даже при использовании больших объёмов памяти для хранения словоформ.

Лекция 11. Обработка цифрового звука

Микширование

Сведéние или *микширование* (*mixing, смешение*) — стадия создания из отдельных записанных треков конечной записи, следующий после звукозаписи этап создания фонограммы, заключающийся в отборе и редактировании (иногда реставрации) исходных записанных треков, объединении их в единый проект и обработке эффектами. Редактирование часто выделяется в самостоятельный этап работы. Стадия сведения исторически возникла в 60-е годы XX века с появлением многодорожечных магнитофонов, позволивших записывать 8 каналов одновременно. Первоначально, сведение не являлось отдельной стадией процесса создания фонограмм, а было составным элементом звукозаписи, так как технические возможности сохранять весь проект отсутствовали, и сведение осуществлялось по мере записи нового материала без возможности возврата к более ранним стадиям. С увеличением числа каналов стала возможной работа со всем проектом, что выделило сведение в самостоятельный этап работы. В 90-е годы XX века значительное распространение получили цифровые технологии сведения. В 1994 году с выпуском программно-аппаратного комплекса Pro Tools III, стала возможной обработка эффектами реального времени, позволившая подбирать параметры обработки во время прослушивания. Cubase VST в 1996 году положил начало сведению в том виде, в котором оно наиболее распространено по настоящее время.

Сведение в проектах электронной музыки — этап следующий после её создания. Этап звукозаписи при работе над электронным проектом чаще всего отсутствует. Граница между созданием и сведением электронной музыки размыта, проект попадает на сведение уже частично сведенным, так как многие виртуальные синтезаторы уже имеют обработки.

В результате сведения многоканальный проект выводится в монофоническую, стереофоническую или многоканальную фонограмму, которая, обычно, получает свой окончательный вид в процессе именуемом мастеринг.

Сведение — не чисто технический процесс соединения различных треков в единое целое, это скорее творческая деятельность, от которой зависят особенности звучания результата. Цель сведения различается в зависимости от концепции проекта.

Критерием оценки в проектах, ориентированных на реалистическое отображение процесса исполнения музыки, служит протокол OIRT (International Radio and Television Organisation):

- *Пространственное впечатление записи*, которое включает в себя такие составляющие, как: передача ощущения объёма помещения, в котором располагаются звуковые источники, естественность передаваемого пространства, реверберационные отражения, планы звуковых источников, учёт традиций звукового решения пространственного образа в музыке различных стилевых направлений, отсутствие дефектов вызванных наложением нескольких звуковых пространств (многопространственность).
- *Прозрачность фонограммы*, которая определяется текстовой разборчивостью, различимостью звучания отдельных инструментов или групп инструментов, ясностью передаваемого пространства.
- *Музыкальный баланс фонограммы* создаваемый логичными соотношениями громкостей между частями произведения, соотношениями громкостей голосов, инструментальных групп и отдельно взятых инструментов.
- *Тембр фонограммы*, как целостного произведения, удобство восприятия тембрового звучания, естественность отражения тембральной окраски инструментов и выгодность подачи тембров.
- *Стереофоничность фонограммы*, как целостной звуковой панорамы, характеризующейся симметричностью положения прямых сигналов и отражений, равномерностью и естественностью расположений звуковых источников, учетом традиций решения панорамирования в музыке различных стилей.
- *Технические замечания по качеству звукового образа*, дефектов, проявляющихся в нелинейных искажениях, неверной передаче частотной характеристики, резонансах, различных видах помех и шумов.
- *Характеристика исполнения*, которая заключается не только в отсутствии исполнительского брака, заключающегося в неверных нотах, ритмических ошибках, интонационных ошибках, недостаточной ансамблевой сыгранности, но и в качестве использования выразительных средств, таких как темп и его агогические отклонения, пропорциональность динамических оттенков динамического плана произведения в целом и градаций динамики на уровне интонации.
- *Аранжировка* произведения представленного в виде фонограммы оценивается в случае исполнения переложений музыкальных произведений для других составов исполнителей.
- *Динамический диапазон фонограммы* в протоколе отражает не только соотношение полезного сигнала и шумов, соотношение уровней звуков между пиками и самыми тихими фрагментами, но и соответствие динамики условиям, в которых фонограмма будет прослушиваться, традиционным для конкретных стилей музыки представлениям о решении динамического плана, естественность и логичность в передаче динамических оттенков, акцентов и кульминаций.

Данная концепция используется при сведении большей части академических жанров, значительной части фольклорной и джазовой музыки. При этом, часто, наилучшие результаты могут быть получены, если условия записи были близки к идеальным, а запись выполняется на стереопару микрофонов и вообще не требует сведения.

При сведении поп-, рок-, электронных и хип-хоп проектов чаще всего фонограмма рассматривается не как документальная фиксация создания музыкантами произведения искусства, а как самостоятельный вид искусства, выражающий свой собственный художественный образ. Для оценки сведения в рамках данной концепции не существует единого протокола.

Мастеринг

Мастеринг звукозаписей — процесс подготовки и переноса записанной и сведенной фонограммы на какой-либо носитель для последующего размножения.

При подготовке к мастерингу полученные после записи исходники редактируются; при необходимости также осуществляется домонтаж/подмонтаж частей произведения. При необходимости могут применяться другие приёмы работы со стереофонограммой: шумоподавление, полосовая компрессия, расширение/сужение стереобазы и т. д. Таким образом и получают материал для создания каждой аудиокomпозиции.

Затем, когда все в отдельности композиции (мастер-треки) готовы, их собирают в нужном порядке, добавляют между ними переходы (напр. добавляются натуральные шумы и т. д.) либо паузы и получают финализированный продукт (аудиоальбом, фонограмму фильма, микстейп и т. д.)

Исходный материал обрабатывается путём эквализации и компрессии (лимитирования). Для каждого вида издания (носителя) обычно применяются собственные настройки мастеринга в зависимости от типа носителя, аудиопотребителя и воспроизводящего звук прибора (особенностей его звучания).

Виртуальный цифровой мастеринг (ремастеринг) — альтернативный к классическому аппаратному мастерингу звуковых фонограмм, в котором вся редакция осуществляется в виртуальной цифровой среде, как правило, внутри компьютерной системы с помощью звуковых редакторов или специального программно-аппаратного комплекса, в котором звук не проходит дополнительных аналогово-цифровых преобразований, оставаясь в цифровом формате. Получил широкое распространение с конца 1990-х годов.

Последовательность мастеринга:

Удаление артефактов. Щелчки, шумы, фон, неправильно поставленная дикция (шипящие и свистящие), западание некоторых частот — все это нам нужно ликвидировать.

Шумоподавление - процесс для удаления шумов готовых фонограмм. Программы или модули (как встроенные, так и внешние), осуществляющие такой тип шумоподавления, называются *денойзерами* (denoiser).

Главным функциональным элементом шумоподавления является *фильтр*, изменяющий полосу пропускания в зависимости от спектра обрабатываемого сигнала. Управляющая этим фильтром электронная схема анализирует входной сигнал, а затем на основе этого анализа меняет параметры фильтра таким образом, чтобы обеспечить максимально благоприятные условия для передачи полезных составляющих сигнала и одновременно максимально ослабить его шумовые компоненты.

Если коснуться практики, то на сегодняшний день большая часть работ по очистке звука сводится к избавлению сэмпла от высокочастотного шума (hiss). Существует два метода. Первый из них - фильтрация. Для начала нужно отфильтровать все верхние частоты, которые забиты именно шумом. А второй метод заключается в том, чтобы программа сканировала образец шума, а затем по алгоритму вычла бы этот шум из волны.

Большинство выпускаемых денойзеров в качестве самого фильтра применяют только фильтр низких частот, который обрезает все высокочастотные помехи, лежащие за пределами полосы частот с полезным сигналом.

У денойзеров помимо достоинств есть также крупный и неустранимый недостаток: вследствие поочередного исчезновения из звукового сигнала отдельных узких полос, его составляющих, сигнал приобретает характерный «ме-таллический» призыв. К сожалению, его невозможно устранить, хотя этот звук можно сделать достаточно малозаметным путем тщательного подбора параметров работы самого денойзера.

Эквализация (Соблюдение частотного баланса). Эквализация фонограммы при мастеринге должна, как правило, носить больше косметический характер. Если исходный материал имеет некоторые артефакты, то его лучше пересвести заново. К таким артефактам чаще всего относятся «инструментальная каша» в диапазоне от 500 Гц до 3 кГц, переизбыток реверберации, противофаза и т.д. Частотная картина композиции должна быть единым целым, где понятно расположение инструментов, слышна их отчетливость и прозрачность.

Динамическая обработка.

Максимайзер (цифровой пиковый лимитер) - это устройство динамической обработки, повышающий уровень звукового сигнала. Громкость звука зависит не только от уровня звука (или звукового давления), но и от его спектрально-временного состава. Если частотный баланс фонограммы уже определен и менять его нежелательно, то для повышения громкости фонограммы нужно повышать уровень сигнала. Зачем повышать громкость? На это есть две причины. Первая из них заключается в том, что громкая музыка чаще всего кажется "красивее", чем тихая, и больше привлекает внимание. Поэтому большинство продюсеров всеми силами стремятся повысить уровень фонограммы при мастеринге: ведь от этого может зависеть ее коммерческий успех. Вторая причина повышения громкости - желание наиболее полно использовать динамический диапазон носителя аудиозаписи, будь то компакт-диск или аналоговая лента. Также важно максимально использовать динамический диапазон воспроизводящего устройства, чтобы запись не тонула в шумах. При записи звука носители обычно ограничивают величиной перегрузки *пиковый* уровень сигнала, а не его среднеквадратичную мощность (это несколько упрощенная, но близкая к реальности модель для большинства аналоговых и цифровых носителей звука). Отношение пикового уровня фонограммы к ее среднеквадратичному уровню (RMS) называется *пик-фактором* (crest-factor, кросс-фактор). Прямоугольная волна (меандр) имеет единичный пик-фактор 0 дБ. Пик-фактор синусоиды равен 3 дБ. Фонограммы с широкой динамикой или резкими пиками обладают высоким пик-фактором (20 дБ и более), а сильно компрессированные фонограммы - низким пик-фактором (10...15 дБ). Ясно, что при ограниченной пиковой мощности, фонограмма с меньшим пик-фактором может достичь более высокой громкости. С целью уменьшения пик-фактора фонограммы ее пропускают через устройства динамической обработки (рис. 1). Рассмотрим их принципы работы.

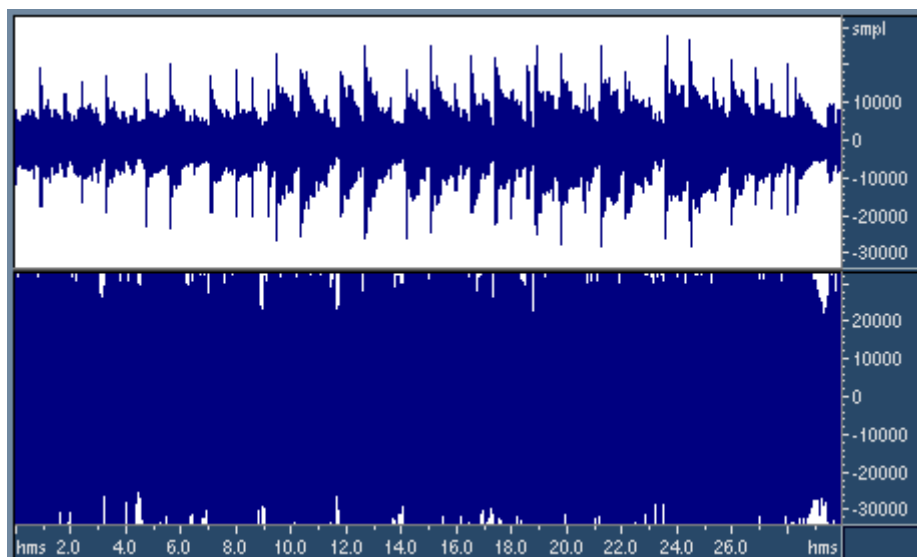


Рис. 1. Фонограмма до динамической обработки и после. Уменьшение пик-фактора. Клиппинга здесь нет, и звук вполне приемлем для радио.

Основными устройствами для работы с уровнями фонограммы являются *устройства динамической обработки*. Принцип действия этих устройств заключается в анализе уровня входящего в них аудио-сигнала и изменении этого уровня по некоторому закону. Основными параметрами устройств динамической обработки являются передаточная характеристика и время атаки/восстановления.

Компрессор (от англ. «compress» — сжимать, сдавливать) — это электронное устройство или компьютерная программа, выполняющее уменьшение (сжатие) динамического диапазона звукового сигнала; иными словами, компрессор позволяет сделать более узкой разницу между самым тихим и самым громким звуком.

Передаточная характеристика (не путать с амплитудно-частотной характеристикой) — это зависимость желаемого выходного уровня звука от входного уровня. В соответствии с передаточной характеристикой, устройство динамической обработки определяет тот коэффициент усиления, который нужно применить к входному сигналу в каждый момент времени. Пример передаточной характеристики показан на рис. 2.

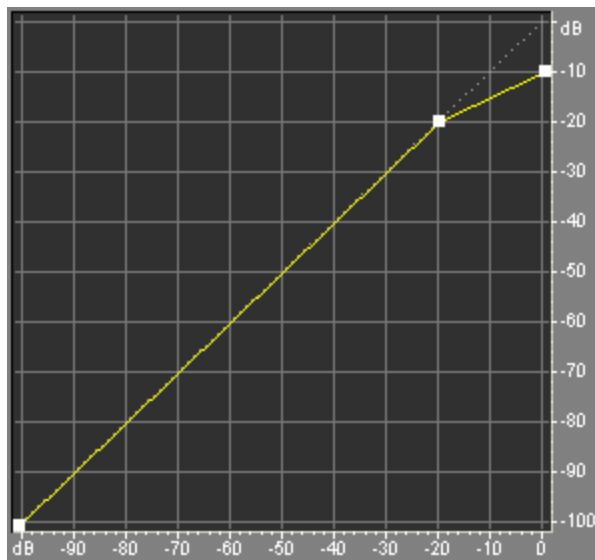


Рис. 2. Передаточная характеристика компрессора. Порог равен -20 дБ, степень компрессии 2:1.

Перелом в передаточной характеристике называется *точкой неергиба* (knee). Входной уровень, соответствующий колену, называется *порогом* (threshold). Угол наклона передаточной характеристики выше порога определяет *степень сжатия* (ratio, степень компрессии). Степень сжатия 2:1 означает, что при увеличении входного уровня на 2 дБ выше порога выходной уровень вырастет лишь на 1 дБ. Если степень сжатия равна единице, то уровень звука при прохождении через прибор не изменится. Если она стремится к бесконечности, то устройство будет ограничивать амплитуду выходного звука значением порога. Такие устройства называются *лимитерами*, они ограничивают динамический диапазон. Если же степень сжатия меньше единицы, например 1:1.5, то это значит, что когда входной уровень превышает порог, устройство будет повышать выходной уровень по сравнению с входным. Такие устройства называются *экспандерами*, они расширяют динамический диапазон. Существуют и другие типы устройств динамической обработки: гейты, дакеры, левелеры и пр., со своими специфическими передаточными характеристиками и параметрами работы.

Иногда передаточная характеристика сглаживается, чтобы в ней не было острых углов

(рис. 3). Этот режим называется *soft knee* или *soft threshold* (мягкий порог). Компрессор с мягким порогом начинает немного уменьшать уровень сигнала еще до достижения им величины порога.

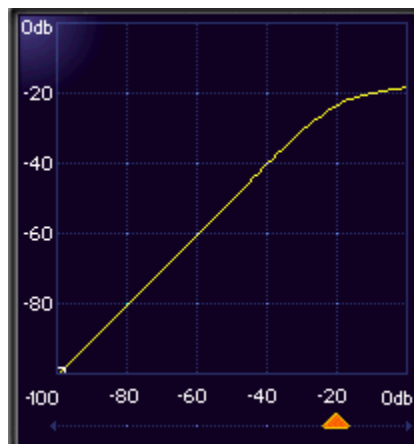


Рис. 3. Мягкий порог.

Работу устройства динамической обработки можно описать следующей схемой. Устройство следит за входным уровнем и в соответствии с ним регулирует выходной уровень, т.е. *применяет к входному сигналу некоторую амплитудную огибающую (коэффициент усиления), меняющуюся со временем.* Для хорошего звучания результирующего сигнала нужно соблюсти несколько условий. Самое главное из них таково: амплитудная огибающая должна быть гладкой, без разрывов и, по возможности, без изломов. Действительно, если амплитудная огибающая имеет разрывы, то выходной звук тоже будет иметь разрывы в форме волны, слышимые как щелчки и треск. Изломы в амплитудной огибающей также будут приводить к искажениям выходного сигнала.

Для сглаживания амплитудной огибающей в устройствах динамической обработки имеются два параметра: *время атаки* (attack, время срабатывания) и *время восстановления* (release). Они определяют скорость реакции устройства на изменения входного уровня. Время атаки показывает, за какое время устройство реагирует на превышение порога (атаку), а время восстановления показывает, за какое время устройство реагирует на возвращение входного уровня обратно под порог.

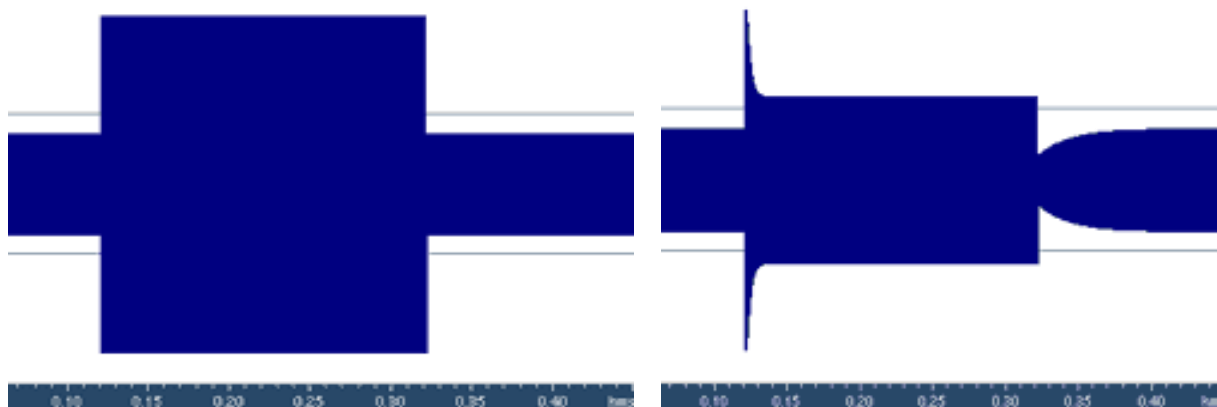


Рис. 4. Звук до и после обработки компрессором.

Пусть на вход компрессора подается сначала слабый сигнал, не превышающий порога, а затем - атака, превышающая порог (рис. 4). В соответствии с передаточной характеристикой, компрессор должен пропустить слабый сигнал без изменения, а уровень громкого сигнала (атаки) - ослабить. Время атаки указывает, за какое время компрессор изменит свой коэффициент усиления от единичного до результирующего, предписанного передаточной характеристикой. Если вслед за громким сигналом входной уровень снова опускается ниже порога, то компрессор переходит в стадию восстановления и снова увеличивает свой коэффициент усиления до единичного. Время, за которое коэффициент усиления вернется к единичному значению, и будет временем восстановления.

Определения времени атаки и восстановления могут различаться у разных производителей. В некоторых устройствах под временем восстановления понимается не полное время возвращения коэффициента усиления, а время его возвращения, скажем, до половины обратного пути. Часто коэффициент усиления возвращается к исходному значению по экспоненте, и в этом случае лишь второе определение имеет смысл. В некоторых устройствах время атаки задается скоростью изменения коэффициента усиления (дБ/сек) или, наоборот, временем изменения коэффициента усиления на 6 дБ.

Время атаки и время восстановления измеряются в миллисекундах и могут меняться в широких пределах для различных приборов динамической обработки и в зависимости от конкретных задач. Так, например, в компрессорах обычное время атаки - порядка

10...100 мс, а типичное время восстановления - порядка 100...1000 мс. Ясно, что чем больше время атаки и восстановления, тем медленнее будет меняться во времени амплитудная огибающая, тем более гладкой она будет. Однако при большом времени атаки компрессор будет пропускать короткие атаки, превышающие порог, т.к. не будет успевать на них реагировать. Это может быть нежелательным, например, для лимитеров.

Создание звуковой картины. В определенных случаях требуется добавление реверберации, расширение стереобазы. При этом вы должны понимать, что внедрение трехмерного звука и реверберации может повлиять в худшую сторону на параметр RMS. Поэтому нужно выбирать между абсолютной прослушиваемостью и эстетическим фактором.

На самом деле многие звукорежиссеры уже не раз отказывались от реверберации в пользу более громкого звучания фонограммы, но тогда очень часто терялся этот самый эстетический фактор. Как вариант — использовать системы многополосной реверберации, где вы можете добавить эффект только для определенного частотного диапазона. Картина получается более аккуратной. В принципе по правилам психоакустики достаточно только обозначить присутствие эффекта, и слушатель будет его ощущать подсознательно. Чтобы получить более яркое звучание, часто необходимо использовать такое устройство, как эксайтер, который позволяет насытить сигнал дополнительными гармониками и улучшить параметры стереозвучания.

Аудиоредакторы

Sony Sound Forge. Программа Sound Forge долгое время являлась флагманским продуктом компании Sonic Foundry, а в мае 2003 года успешно переключалась под другую торговую марку — Sony Pictures Digital. К радости огромной армии поклонников этого звукового редактора, подразделение Sony продолжило развитие Sound Forge. На сегодняшний день это отличный и удобный цифровой аудиоредактор, в состав которого входит набор утилит, предназначенных для работы со звуком. При помощи данной программы можно эффективно обрабатывать звуковые композиции, накладывать множество эффектов, осуществлять точную и быструю запись и редактирование сэмплов, кодировать данные, конвертировать звуковые файлы в различные форматы и т.д.

Это одна из самых популярных и действительно полезных программ, предназначенных для обработки аудио на профессиональном уровне. С ее помощью можно менять аудиосигнал до невозможности или же редактировать записанную партию любого музыкального инструмента. Sony Sound Forge успешно объединяет в себе практически полный набор современных звуковых эффектов и мощные средства редактирования звуков для последующего их использования в сэмплере.

Отличительные особенности Sound Forge :

- функциональность и эргономичность интерфейса;
- большое количество возможностей редактирования и обработки эффектами;
- поддержка большинства новых технологических стандартов, включая аудио с характеристиками 32/64 бита с плавающей точкой и частотой дискретизации от 2 до 192 кГц;
- постоянная поддержка со стороны разработчика.

Adobe Audition. Эта программа имеет достаточно длинную историю. Когда-то она носила название Cool Edit и была довольно упрощенным вариантом нынешней программы. По существу, первая версия Audition являла собой клон последней версии Cool Edit Pro (на которой и закончилась история этого редактора) На сегодняшний день Adobe Audition — профессиональный инструмент для микширования и обработки аудиоматериалов.

Программа предоставляет средства для микширования, редактирования, создания мастер-копий и обработки аудиоэффектов:

- редактирование отдельных аудиофайлов;
- смешивание одновременно до 128 треков;
- создание петель и применение более 45 цифровых встроенных эффектов;
- производство высококачественной аудиопродукции: Adobe Audition поддерживает файлы с разрядностью до 32 бит и частотой дискретизации до 192 кГц;
- поддерживается запись на любые носители (включая магнитную пленку, диски CD, DVD или DVD- Audio). Вся внутренняя обработка выполняется с 32-битной разрядностью, что позволяет достичь высокого качества выходного материала;
- удобный интерфейс;
- встроенные режимы просмотра (многодорожечный и для редактирования);
- спецэффекты реального времени;
- средства анализа;
- функция восстановления звука и поддержка видеоряда;
- поддержка виртуальных инструментов VSTi;
- современные средства микширования;
- средства для адаптивного устранения шума и инструменты для поэтапной коррекции, которые помогают восстановить старые записи.

Audacity. Audacity — это небольшой кроссплатформенный бесплатный аудиоредактор, работающий под Windows, Linux/ Unix, Mac OS 9 и Mac OS X. Он создан группой энтузиастов на C/C++ и распространяется под лицензией GPL.

Вполне приличное количество встроенных средств обработки звука: усиление (Amplify), усиление басов (Bass Boost), эхо, вау-вау (Wah-wah), все виды фейдов, удаление шума и т.д. Кроме того, версии под Windows и Mac OS знают VST-плагины. Linux-версия Audacity поддерживает и вовсе несметное количество различных плагинов — уже больше сотни.

В Audacity есть стандартный набор встроенных эффектов и фильтров: шумоподавление, усиление/ослабление сигнала, фэйзер, эхо, FFT-фильтр и др. В комплект Audacity также входит бесплатный VST-ревербератор Freeverb2. Кроме того, вы можете создавать собственные плагины на основе скриптового языка, поддерживаемого Audacity. Другими словами, у вас есть возможность не только написать математическую формулу для обработки звукового сигнала, но и сформировать под это дело графический интерфейс. Такие плагины просто появляются в меню Effect, как и обычные эффекты и плагины. Дополнительные плагины можно скачать с сайта редактора.

К сильным сторонам Audacity относятся широкие возможности по обработке сигнала и способность программы узнавать любые форматы аудио — как сжатые (MP3, OGG), так и не сжатые (WAV, AIF).

Благодаря оригинальной системе хранения аудиоданных Audacity очень быстро открывает и записывает свои проект-файлы. Пока работа над аудиосигналом производится непосредственно в самом пакете, сигнал хранится на диске в виде небольших файлов собственного формата Audacity, каждый из которых представляет собой отдельный фрагмент аудио. В Sound Forge и Audition обработка производится непосредственно над цельным wav-файлом, из-за чего запись и открытие файла происходят достаточно медленно. Собственный формат Audacity не совместим ни с какими другими программами, да это и не нужно, учитывая его сугубо служебные функции. Audacity позволяет экспортировать аудио в любой распространенный формат, кроме MP3. Алгоритм кодирования в MP3 защищен копирайтом, так что кодек придется

скачивать отдельно. На сайте проекта Audacity есть нужные ссылки.

Инструментов для редактирования в Audacity не очень много. Это обычные утилиты для нарезки звука, масштабирования и перемещения. Есть еще инструмент для редактирования огибающих громкости.

Не обладая всеядностью некоторых звуковых редакторов, Audacity тем не менее поддерживает чтение и запись таких форматов, как .wav, .aiff, .au, OggVorbis и .mp3. Запись в последний — только при наличии внешнего кодера.