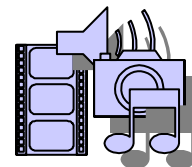


# Лекция 1. Понятие мультимедиа технологии



## 1.1. Составляющие части мультимедиа

Слово «мультимедиа» прочно вошло в наш лексикон, и без него уже трудно представить современный компьютерный мир. Как и всякое удачное понятие, оно многообразно. Пожалуй, наиболее точная формулировка принадлежит одному из пионеров мультимедиа в нашей стране Сергею Новосельцеву: «Мультимедиа (англ. multimedia от лат. multum – много и media, medium – средоточие, средства) – это комплекс аппаратных и программных средств, позволяющих пользователю работать в диалоговом режиме с разнородными данными (графикой, текстом, звуком, видео и анимацией), организованными в виде единой информационной среды».

Как видим, мультимедиа объединяет несколько типов разнородных данных (текст, звук, видео, графическое изображение и анимацию) в единое целое. И это понятие само по себе имеет три лица.

Во-первых, мультимедиа – как идея, т. е. новый подход к хранению информации различного типа в единой цифровой форме.

Во-вторых, мультимедиа – как оборудование для обработки и хранения информации, без него мультимедиа-идею реализовать невозможно.

В-третьих, это программное обеспечение, позволяющее объединить четыре элемента информации в законченное мультимедиа-приложение.

Мультимедиа технологии являются одним из наиболее перспективных и популярных направлений информатики. Они имеют целью создание продукта, содержащего «коллекции изображений, текстов и данных, сопровождающихся звуком, видео, анимацией и другими визуальными эффектами (Simulation), включающего интерактивный интерфейс и другие механизмы управления». Данное определение сформулировано в 1988 году крупнейшей Европейской Комиссией, занимающейся проблемами внедрения и использования новых технологий. Интерактивность – свойство реагировать на действия пользователей, в том числе и управлять пользователем.

Идейной предпосылкой возникновения технологии мультимедиа считают концепцию организации памяти «MEMEX», предложенную еще в 1945 году американским ученым Ваннивером Бушем. Она предусматривала поиск информации в соответствии с ее смысловым содержанием, а не по формальным признакам (по порядку номеров, индексов или по алфавиту и т. п.). Эта идея нашла свое выражение и компьютерную реализацию сначала в виде системы гипертекста (система работы с комбинациями текстовых материалов), а затем и гипермедиа (система, работающая с комбинацией графики, звука, видео и анимации), и, наконец, в мультимедиа, соединившей в себе обе эти системы.

Однако всплеск интереса в конце 80-х годов XX в. к применению мультимедиа технологии в гуманитарной области (и, в частности, в историко-культурной) связан, несомненно, с именем выдающегося американского компьютерщика-бизнесмена Билла Гейтса, которому принадлежит идея создания и успешной реализации на практике мультимедийного продукта на основе служебной музейной инвентарной базы данных с использованием в нем всех возможных «сред»: изображений, звука, анимации, гипертекстовой системы («National Art Gallery London»).

Именно этот продукт аккумулировал в себе три основных принципа мультимедиа:

- представление информации с помощью комбинации множества воспринимаемых человеком сред;
- наличие нескольких сюжетных линий в содержании продукта (в том числе и выстраиваемых самим пользователем на основе «свободного поиска» в рамках предложенной в содержании продукта информации);
- художественный дизайн интерфейса и средств навигации.

Несомненным достоинством и особенностью технологии являются следующие возможности мультимедиа, которые активно используются в представлении информации:

- возможность хранения большого объема самой разной информации на одном носителе;
- возможность увеличения (детализации) на экране изображения или его наиболее интересных фрагментов, иногда в двадцатикратном увеличении (режим «лупа») при сохранении качества изображения. Это особенно важно для презентации произведений искусства и уникальных исторических документов;
- возможность сравнения изображения и обработки его разнообразными программными средствами с научно-исследовательскими или познавательными целями;
- возможность выделения в сопровождающем текстовом или другом визуальном материале «горячих слов (областей)», по которым осуществляется немедленное получение справочной или любой другой пояснительной (в том числе визуальной) информации (технологии гипертекста и гипермедиа);
- возможность осуществления непрерывного музыкального или любого другого аудиосопровождения, соответствующего статичному или динамичному визуальному ряду;
- возможность использования видеофрагментов из фильмов, видеозаписей и т. д., функции «стоп-кадра», покадрового «пролистывания» видеозаписи;
- возможность включения в содержание диска баз данных, методик обработ-

ки образов, анимации (к примеру, сопровождение рассказа о композиции картины графической анимационной демонстрацией геометрических построений ее композиции) и т. д.;

- возможность подключения к глобальной сети Internet;
- возможность работы с различными приложениями (текстовыми, графическими и звуковыми редакторами, картографической информацией);
- возможность создания собственных «галерей» (выборки) из представляемой в продукте информации (режим «карман» или «мои пометки»);
- возможность «запоминания пройденного пути» и создания «закладок» на заинтересовавшей экранной «странице»;
- возможность автоматического просмотра всего содержания продукта («слайд-шоу») или создания анимированного и озвученного «путеводителя-гида» по продукту («говорящей и показывающей инструкции пользователя»), включение в состав продукта игровых компонентов с информационными составляющими;
- возможность «свободной» навигации по информации и выхода в основное меню (укрупненное содержание), на полное оглавление или вовсе из программы в любой точке продукта.

Итак, мультимедийный продукт – наиболее эффективная форма подачи информации в среде компьютерных информационных технологий. Он позволяет собрать воедино огромные и разрозненные объемы информации, дает возможность с помощью интерактивного взаимодействия выбирать интересующие в данный момент информационные блоки, значительно повышая эффективность восприятия информации.

## **1.2. Классификация мультимедиа приложений**

Мультимедиа – это взаимодействие визуальных и аудиоэффектов под управлением интерактивного программного обеспечения. Мультимедиа – комбинация текста, графических изображений, звука, анимации и видео-элементов.

Согласно представленным выше определениям, мультимедиа можно классифицировать с разных точек зрения:

- на основе поддержки взаимодействия,
- на основе использования различных мультимедийных телекоммуникационных технологий.

## **1.3. Области применения мультимедиа приложений**

Области применения:

- Обучение с использованием компьютерных технологий (научно-просветительская или образовательная сфера);

- Видеоэнциклопедии, интерактивные путеводители, тренажеры, ситуационно-ролевые игры и др.;
- Информационная и рекламная служба;
- Популяризаторская и развлекательная сферы;
- Интернет-вещание;
- Развлечения, игры, системы виртуальной реальности;
- Презентационная (витринной рекламы), СМИ;
- Творчество (станция мультимедиа становится незаменимым авторским инструментом в кино и видеоискусстве. Автор фильма за экраном такой настольной системы собирает, «аранжирует», создает произведения из заранее подготовленных – нарисованных, отснятых, записанных и т. п. – фрагментов;
- Военные технологии;
- Промышленность и техника (сенсорные экраны);
- Торговля.
- В научно-исследовательской области – это электронные архивы и библиотеки – для документирования коллекций источников и экспонатов, их каталогизации и научного описания, для создания «страховых копий», автоматизации поиска и хранения, для хранения данных о местонахождении источников для хранения справочной информации, для обеспечения доступа к внемузейным базам данных, для организации работы ученых не с самими документами, а с их электронными копиями и т. д.).
- Медицина: базы знаний, методики операций, каталоги лекарств и т. п.
- Искусственный интеллект – внедрение элементов искусственного интеллекта в системе мультимедиа. Они обладают способностью «чувствовать» среду общения, адаптироваться к ней и оптимизировать процесс общения с пользователем: они подстраиваются под читателей, анализируют круг их интересов, помнят вопросы, вызывающие затруднения и могут сами предложить дополнительную или разъясняющую информацию.
- Системы распознавания речи, понимающие естественный язык, еще более расширяют диапазон взаимодействия с компьютером.

#### **1.4. Аппаратные средства мультимедиа технологии**

Технологию мультимедиа составляют специальные аппаратные и программные средства.

Для построения мультимедиа системы необходима дополнительная аппаратная поддержка: аналого-цифровые и цифроаналоговые преобразователи для перевода аналоговых аудио- и видеосигналов в цифровой эквивалент и обратно, видеопроцессоры для преобразования обычных телевизионных сигналов к виду, воспроизводимому электронно-лучевой трубкой

дисплея, декодеры для взаимного преобразования телевизионных стандартов, специальные интегральные схемы для сжатия данных в файлы допустимых размеров и т. д.

### **1.4.1. Аппаратные средства**

Все оборудование, отвечающее за звук, объединяется в звуковые карты, а за видео – в видеокарты.

Аппаратные средства мультимедиа:

- Средства звукозаписи (звуковые платы, микрофоны);
- Средства звуковоспроизведения (усилитель, колонки, акустические системы, наушники и гарнитуры);
- Манипуляторы (компьютерные мыши, джойстики, миди-клавиатуры);
- Средства «виртуальной реальности» (перчатки, очки, шлемы виртуальной реальности, используемые в играх);
- Носители информации (CD, DVD и HDD);
- Средства передачи (мини видеокамеры, цифровые фотоаппараты);
- Средства записи (приводы CD / DVD-ROM , CDRW / DVD+RW, TV- и FM-тюнеры);
- Средства обработки изображения (платы видеомонтажа, клавиатуры, графические акселераторы).
- Компьютер, телевизор, средства для получения и удобного восприятия информации и др.

## **1.5. Программные средства мультимедиа технологии**

Программные средства мультимедиа складываются из трех компонентов:

1. Системные программные средства.
2. Инструментальные программные средства.
3. Прикладные программные средства.

### **1.5.1. Системные программные средства**

Системные программные средства – это набор программ, входящих в состав операционной системы компьютера и осуществляющих управление устройствами мультимедиа, причем это управление на двух уровнях – физическое управление вводом-выводом информации на низком уровне с помощью машинных команд и управление пользователем характеристиками устройств с помощью графического интерфейса, изображающего пульт управления устройством, например регулировки громкости звука, тембра, стереобаланса и т. д. Как правило, программы физического управления устройствами называют драйверами устройств.

### **1.5.2. Инструментальные программные средства**

Инструментальные программные средства – программы позволяющие модифицировать мультимедийные файлы и создавать мультимедийные приложения.

Инструментальные программные средства – это пакеты программ для создания мультимедийных приложений:

- редакторы неподвижных графических изображений,
- средства создания анимированных GIF-файлов,
- средства аудио- и видеомонтажа,
- средства создания презентаций,
- средства распознавания текстов, введенных со сканера,
- средства создания обучающих программ,
- системы распознавания голоса и преобразования звуковых файлов в текстовые,
- системы создания приложений виртуальной реальности и другие.

Инструментальные средства существенно расширяют возможности управления мультимедийными устройствами по сравнению с теми, которые предоставляют системные средства, но это всегда платные продукты и некоторые из них стоят очень дорого, например профессиональные системы видеомонтажа.

### **1.5.3. Прикладные программные средства**

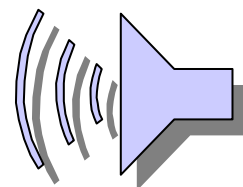
Прикладные программные средства – это готовые и, как правило, продаваемые программные системы на CD или DVD дисках – фильмы, учебники, энциклопедии, игры, книги, виртуальные музеи, путеводители, рекламные материалы и т. д.

### **Вопросы по лекции**

1. Понятие мультимедиа технологии.
2. Каковы аппаратные средства использования мультимедиа технологии?
3. Какими могут быть области применения мультимедиа приложений?
4. Что было идейной предпосылкой возникновения технологии мультимедиа?
5. Дайте определение гипертекста.
6. Какие на современном этапе существуют инструментальные средства для создания гипертекста?
7. Какие в настоящее время есть инструментальные средства для использования гипертекста?
8. Назовите три составляющих мультимедиа.

## Лекция 2. РАБОТА СО ЗВУКОМ.

### ОСНОВНЫЕ СВОЙСТВА СЛУХА



#### 2.1. Введение

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

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

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

**В цифровой звукотехнике** до сих пор существует необходимость понижения шума квантования, возникающего при аналого-цифровом преобразовании. Этот шум действует на слух значительно более раздражающе, чем шум аналоговой аудиоаппаратуры. Хотя максимальное отношение сигнал-шум для цифровой аудиоаппаратуры достигает 96–120 дБ, при воспроизведении записи реальных музыкальных программ оно не превышает 50–70 дБ. Существенное снижение шума квантования при разработке цифровой аудиоаппаратуры также возможно только с учетом особенностей слуха человека.

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

**В области цифровой звукозаписи** для повышения качества воспроизведения звука требуется повышение частоты дискретизации и числа

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

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

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

## 2.2. Восприятие по частоте

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

Прежде всего, необходимо отметить, что слуховая система способна различать высоту звука только у периодических сигналов, поэтому определяющим параметром для различения высоты тона является частота сигнала. Если это сложный звук, то высоту слуховая система может присво-



ить по его основному тону, но только если он имеет периодическую структуру, т. е. его спектр состоит из гармоник (обертонов, частоты которых находятся в целочисленных отношениях). Если это условие не выполняется, то определить высоту тона слуховая система не может. Например, звуки таких инструментов как тарелки, гонги и др. не имеют определенной высоты.

Изучение связи частоты звука и воспринимаемой высоты предпринималось еще Пифагором, а также многими известными физиками: Галилеем, Гельмгольцем, Омом и др. В настоящее время на основе тщательных экспериментов, в процессе которых слушателю предъявлялись два звука разной частоты с просьбой расположить их по высоте, установлена зависимость высоты тона от частоты сигнала, показанная на рис. 2.1.

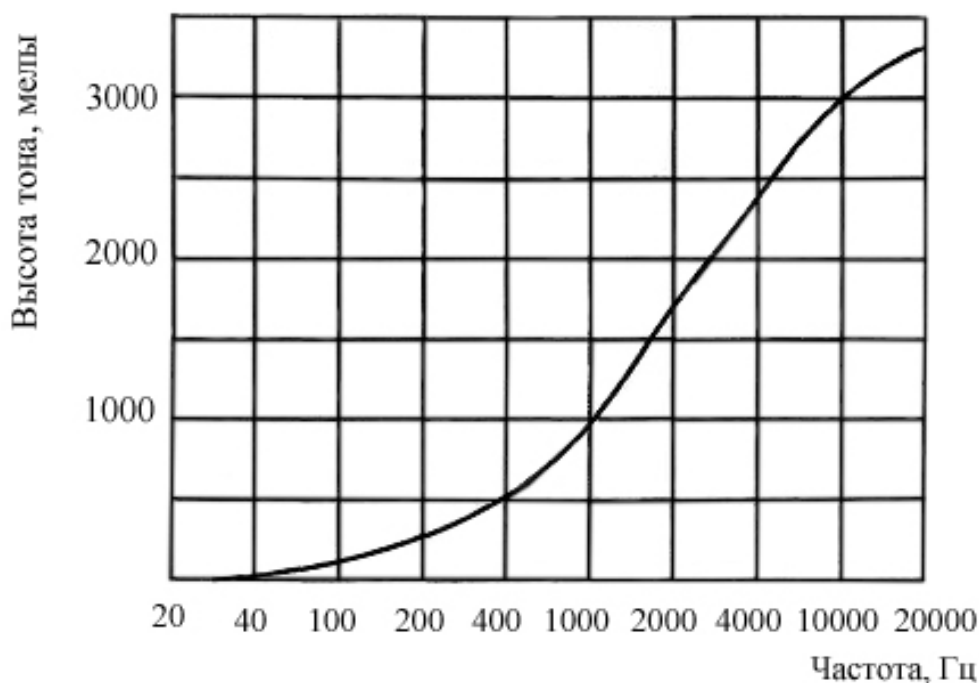


Рис. 2.1. Связь частоты и высоты тона

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

Ощущение высоты чистого тона связано не только с частотой, но и с интенсивностью звука и его длительностью. Как показали различные исследования, при повышении интенсивности звука громкие низкие звуки кажутся еще ниже, а высокие звуки с повышением громкости кажутся

слегка выше (только для средних частот 1–2 кГц влияние интенсивности незаметно). Следует отметить, что эта зависимость весьма незначительна, а для сложных музыкальных звуков почти незаметна.

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

### 2.3. Порог слышимости и болевой порог

Если волокно основной мембраны при своих колебаниях не достает до ближайшей волосковой клетки, то человек не воспринимает звук. При увеличении амплитуды колебаний волокна, как только оно коснется ближайшей клетки, произойдет раздражение нервного окончания, которое сразу же начнет посылать электрические импульсы в слуховой центр мозга; звук будет услышан.

Для того чтобы в полной тишине звук с частотой 1000 Гц был услышан, амплитуда давления вблизи человеческого уха должна достигать всего лишь  $2,84 \cdot 10^{-5}$  Н/м<sup>2</sup> (эффективное значение –  $2 \cdot 10^{-5}$  Н/м<sup>2</sup>), что составляет всего лишь  $2 \cdot 10^{-10}$  атмосферного давления. Интенсивность соответствующей плоской волны в воздухе при этом составит  $10^{-12}$  Вт/м<sup>2</sup>. Интересно заметить, что амплитуда смещения частиц воздуха при этом меньше десятой доли радиуса молекулы. Величина случайных флуктуаций силы давления на барабанную перепонку, связанная с тепловым молекулярным движением, всего в 5–10 раз меньше силы давления звука, заметного в полной тишине.

Величина звукового давления, которая едва заметна на слух при отсутствии всяких других мешающих шумов и звуков, называется пороговой величиной звукового давления, или, сокращенно, **порогом слышимости**. Определение порога слышимости предпринималось исследователями неоднократно. Было выяснено, что пороги слышимости, определенные у ряда людей, могут сильно различаться. Эти различия имеют случайный характер для группы людей одинакового возраста, имеющих нормальный здоровый слуховой орган. Порог слышимости может изменяться у одного и того же человека в зависимости от состояния организма в данный момент: возбуждения, утомления и т. п. Поэтому надежные сведения о пороге слышимости можно получить только статистическим путем, измерив его в определенных условиях у большого числа людей.

Такие статистические исследования проводились в США (1938–1939 гг.), в Англии (1956–1957 гг.), СССР (1958 г.). На основании международного соглашения в качестве стандарта принята кривая зависимости порога слышимости от частоты для чистого синусоидального звука, приведенная на рис. 2.2. При этом в качестве испытуемых при-

влекались лица в возрасте от 18 до 23 лет с заведомо здоровыми органами слуха.

Как видно из рис. 2.2, порог слышимости сильно зависит от частоты. Звуки в области частот от 2000 до 4000 Гц замечаются при звуковом давлении даже меньшем  $2 \cdot 10^{-5}$  Па. В то же время на нижних и верхних звуковых частотах порог слышимости существенно возрастает. Как бы мы ни увеличивали интенсивность звука на частотах выше 20000 Гц, ощущение звука не возникает, т. е. звуки с частотой выше 20000 Гц лежат за пределами частот слышимых звуков большинства людей. Точно так же дело обстоит со звуками с частотой ниже 16–20 Гц. Частоты ниже 16 Гц называются инфразвуковыми, а выше 20000 Гц – ультразвуковыми.

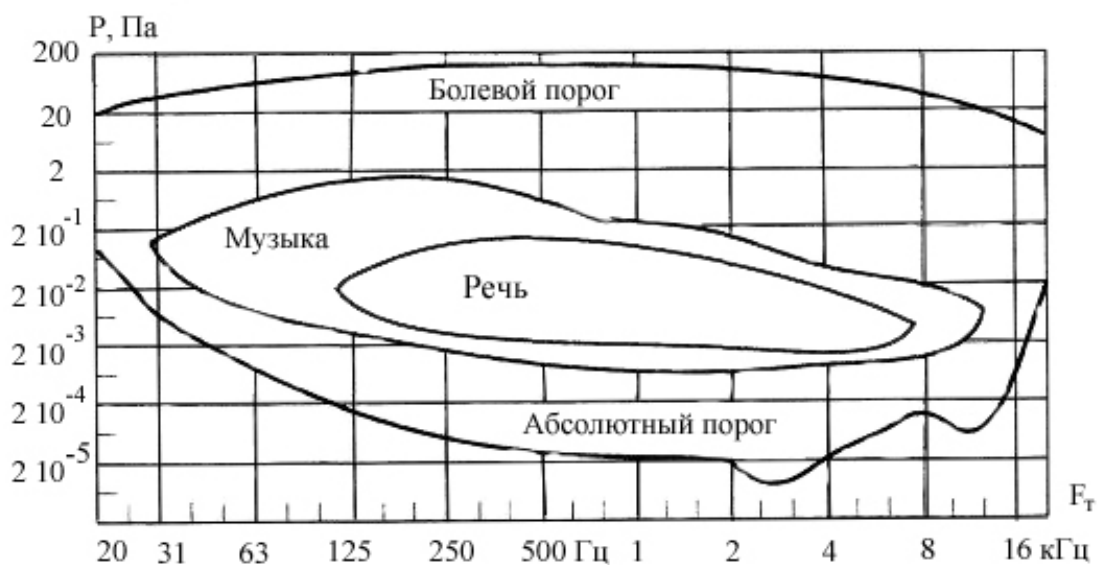


Рис. 2.2. Области слышимости звука

## 2.4. Дифференциальный порог восприятия интенсивности звука

Минимальная различимая на слух разность интенсивности двух звуков одной и той же частоты определяет так называемый **дифференциальный порог слышимости по интенсивности звука**.

Однако сила звука, выраженная в виде звукового давления или потока звуковой энергии, не является мерой величины ощущения или субъективной силы звука, которую принято называть ощущением громкости.

Каким же образом можно оценить громкость как меру субъективной силы звука? Базой для этого может служить психофизический закон, установленный Вебером в 1846 году и заключающийся в том, что **отношение минимального воспринимаемого прироста некоторого стимула к первоначальной его величине является постоянным**. Обозначая силу звука (звукового стимула) через  $J$ , напишем закон Вебера в такой форме:

$$\frac{\Delta J}{J} = \text{const}, \quad (2.1)$$

где  $\Delta J$  – прирост интенсивности, который следует дать звуку интенсивности  $J$ , чтобы стала заметна разница в громкости между звуками  $J$  и  $J + \Delta J$ . Величина  $\Delta J / J$  составляет примерно 10 %. Подчеркнем еще раз, что это соотношение связано со свойствами нервной системы и наблюдается не только при звуковых раздражителях, но и при зрительных, осязательных и др., поэтому оно и носит название всеобщего физиологического закона.

Следующий шаг в развитии теории Вебера был сделан в 1860 году Фехнером, который подверг соотношение (2.1) некоторой математической обработке. Фехнер принял, что прирост  $\Delta J$  есть бесконечно малая величина  $dJ$ , а правую часть считал пропорциональной бесконечно малому приросту ощущения  $dE$  (в данном случае – ощущение громкости). Получается соотношение

$$A \frac{dJ}{J} = dE, \quad (2.2)$$

где  $dJ$  – величина прироста интенсивности (принятая за бесконечно малую);

$dE$  – соответственно «бесконечно малый прирост величины нашего ощущения»;

$A$  – некоторая произвольная постоянная, зависящая от выбора единиц ощущения.

После того, как закон Вебера записан в таком математическом виде, он представляет собой простейшее дифференциальное уравнение, которое можно интегрировать. Интегрируя выражение (1.2), получим:

$$E = A \ln J + C, \quad (2.3)$$

где  $C$  – постоянная интегрирования.

Считая, что на пороге слышимости  $E = 0$  и  $J = J_0$ , мы определим

$$C = -A \ln J_0, \quad (2.4)$$

и получим известный логарифмический закон, носящий название закона Вебера-Фехнера, согласно которому ощущение ( $E$ ) пропорционально логарифму раздражения

$$E = A \ln J + C. \quad (2.5)$$

Переходя к десятичным логарифмам, получим

$$E = 2,3 A \lg \frac{J}{J_0}. \quad (2.6)$$

Константу  $A$  можно принять равной некоторой удобной для вычисления величине, так как она определяет масштаб ощущения и поэтому может быть выбрана произвольно. Принимая, например, условно  $2,3A = 10$ ,

получим

$$E = 10 \lg \frac{J}{J_0}. \quad (2.7)$$

Эта формула дает величину, которую называют уровнем ощущения. Применение шкалы децибел удобно для оценки величины слухового ощущения еще и потому, что минимально ощутимые приросты имеют величину порядка 0,5 дБ.

Вследствие логарифмического закона восприятия и широкого диапазона интенсивностей слышимых звуков для объективной оценки введено понятие **уровня интенсивности**

$$N_I = 10 \lg \frac{I}{I_0}. \quad (2.8)$$

За нулевой уровень условились принимать интенсивность  $I_0$ , близкую к пороговой интенсивности для нормального слуха на частоте 1000 Гц. Эта интенсивность, как уже отмечалось, приблизительно равна  $10^{-12}$  Вт/м<sup>2</sup>. Следовательно, болевой порог интенсивности находится около 120 дБ.

В соответствии с квадратичным соотношением между интенсивностью звука и звуковым давлением уровень звукового давления

$$N_p = 20 \lg \frac{p}{p_0}, \quad (2.9)$$

где  $p_0$  – звуковое давление на нулевом уровне, равное  $2 \cdot 10^{-5}$  Па.

Взяв  $p_0 = 2 \cdot 10^{-5}$  Н/м<sup>2</sup>, что соответствует стандартному порогу слышимости уха человека на частоте 1000 Гц, получим:

$$N_p = 20 \lg p + 94, \text{ дБ}. \quad (2.10)$$

Определенный по (2.9) или (2.10) уровень носит название **уровня звукового давления**, выраженного в децибелах. Численно такое же значение будет иметь и уровень интенсивности звука, определенный по формуле  $N = 10 \lg I + 120$ , дБ, где за  $I$  принята величина пороговой интенсивности  $10^{-12}$  Вт/м<sup>2</sup>.

Уровень звукового давления совпадает по величине с уровнем ощущения только для чистого тона 1000 Гц, так как при этом значение  $I_0$  совпадает с порогом слышимости на 1000 Гц.

Понятие уровня используется не только в акустике, но и электротехнике и связи. Так, под электрическим уровнем понимается величина

$$N = 20 \lg \frac{U}{U_0}.$$

Нулевые уровни электрических величин выбираются при этом так, чтобы мощность, выделяемая при напряжении  $U_0$  на сопротивлении  $R = 600$  Ом, составляла 1 мВт. Отсюда легко найти, что  $U_0 = 0,775$  В.

В телефонии и технике дальней связи многие годы использовалась другая логарифмическая шкала – шкала неперов. Один непер соответствует отношению напряжений, равному 2,718, т. е. основанию натуральных логарифмов.

Таким образом, если  $U/U_0 = e$ , то  $N_{НП} = \text{lm}(U/U_0) = 1$  Нп. Из определения непера следует, что 1 Нп = 8,68 дБ, а 1 дБ = 0,115 Нп.

## 2.5. Пороги слышимости при маскировке

При одновременном воздействии на слух двух звуков один из них может не прослушиваться на фоне другого звука. Этот эффект называется маскировкой. Маскирующее действие различных звуков выявляется путем определения повышения уровня порога слышимости испытательных тонов или узкополосных шумов  $N_{nc}$  относительно уровня абсолютного порога слышимости  $N_{АПС}$ .

Коэффициент маскировки  $K_m$  рассчитывается в децибелах как разность этих двух уровней

$$K_m = N_{ПС} - N_{АПС}$$

**Маскировка чистым тоном.** На рис. 2.3 приведены кривые маскировки тона в диапазоне частот, когда мешающим звуком является тот же тон, но с фиксированной частотой  $F_M = 1$  кГц и разными уровнями звукового давления  $N_M$ . По этим кривым определяется порог слышимости на конкретных частотах маскируемого тона  $E_p$  и рассчитывается коэффициент маскировки  $N_{nc}$ , дБ.

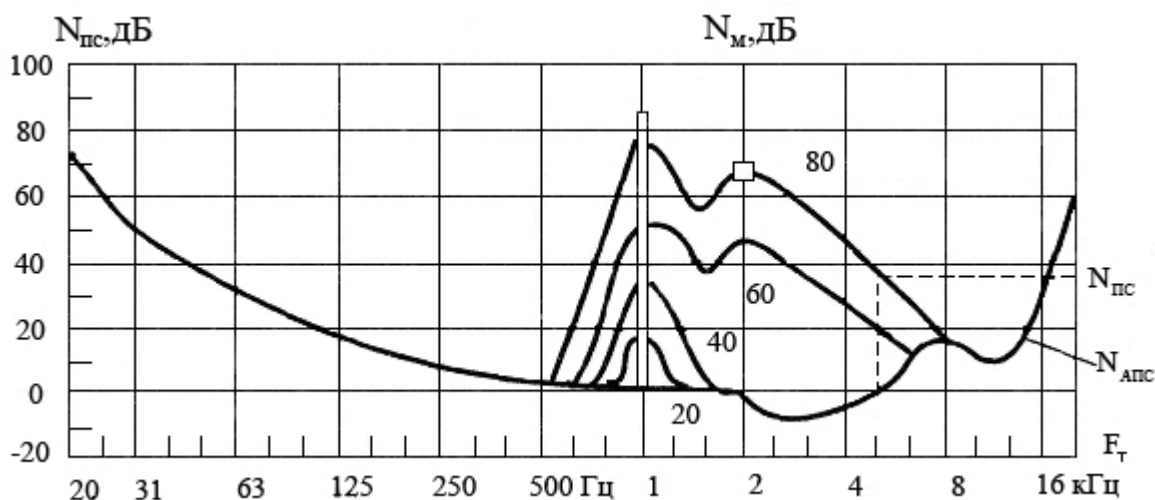


Рис. 2.3. Кривые порога слышимости тона с частотой  $F_T$  при маскировке тоном с частотой  $F_M = 1$  кГц и уровнем  $N_m$ , дБ

Особенности маскировки чистым тоном. Эти особенности могут быть сформулированы следующим образом:

- маскировка максимальна, когда частоты  $F_T$  и  $F_M$  близки, чем дальше по час-

- тоте они расположены друг от друга, тем меньше величина маскировки;
- величина маскировки увеличивается по мере возрастания уровня мешающего тона  $N_M$ ;
  - кривые маскировки существенно несимметричны по шкале частот, у них более резкий спад в сторону низких частот и плавное понижение в сторону верхних частот, поэтому маскировка сильнее проявляется в области высоких частот;
  - сдвиг маскирующего звука по частоте приводит к изменению формы кривой маскировки.

Если частота тона  $F_T$  равна удвоенной, утроенной и выше частоте мешающего тона  $F_M$ , то в широком диапазоне слышимости возникают биения, которые проявляются в виде провалов на графиках кривых маскировки. Эти биения, возникают из-за нелинейности слуха и тем заметнее, чем больше уровень мешающего тона.

## 2.6. Уровень громкости и громкость

### 2.6.1. Уровень громкости

Звуковое давление и интенсивность тонального звука, измеренные приборами, не дают представления об ощущении громкости. Между тем человек может довольно точно установить равенство громкостей двух звуков любого частотного состава. Это свойство слуха было использовано для введения понятия уровня громкости, позволяющего определить громкость одного звука относительно другого, принятого в качестве опорного.

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

Определение уровня громкости тонов других частот производится по графику кривых равного уровня громкости. Эти графики называются изофонами (рис. 2.4). Каждая кривая получена путем субъективного сравнения громкости синусоидального звука произвольной частоты с громкостью звука с частотой 1 кГц. Если звук с частотой 100 Гц имеет уровень 50 дБ, а равногромкий ему звук с частотой 1 кГц имеет уровень 40 дБ, то уровень громкости звука с частотой 100 Гц равен 40 фон.

Для всех изофон характерно небольшое снижение в области частот 3–4 кГц и равномерный подъем при понижении частоты ниже 500 Гц. Величина подъема на разных уровнях неодинакова. При повышении уровня изофоны спрямляются, и подъем становится менее крутым. Это значит, что частотная характеристика слуха при больших уровнях звукового давления более равномерна, чем при малых. При малых уровнях громкости

форма изофон приближается к кривой абсолютного порога слышимости, которой соответствует уровень громкости 3 фона.

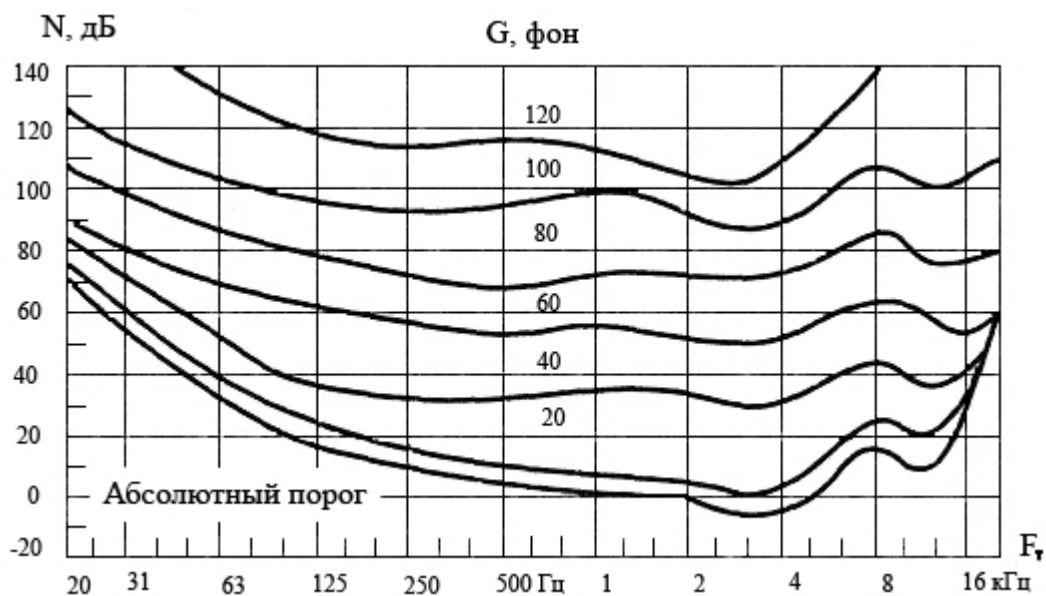


Рис. 2.4. Кривые равного уровня громкости

Чтобы получить представление об уровне громкости звуков, с которыми мы постоянно сталкиваемся в жизни, приведем ряд примеров. Громкий разговор на расстоянии 1 метр – 65–70 фон, уровень громкости в поезде метро – 85–90 фон, обычный разговор – 55–60 фон, уровень громкости в читальном зале библиотеки – 25–30 фон.

Из кривых равного уровня громкости следует, что сохранение естественного звучания, которое соответствует условиям записи, возможно только при условии воспроизведения записи с тем же уровнем.

## 2.6.2. Громкость

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

Необходимость введения понятия громкости связана с тем, что на практике часто представляет интерес не уровень громкости, выраженный в фонах, а величина, показывающая во сколько раз данный звук громче другого. Представляет интерес также вопрос о том, как складываются громкости двух разных тонов. Так, если имеются два тона разных частот с уровнем 70 фон, это не значит, что суммарный уровень громкости равен 140 фон.



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

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

График, показанный на рис. 2.5, позволяет вычислить уровень громкости сложного звука с дискретным спектром. Пусть мы имеем 3 тона: 100, 500 и 3500 Гц с уровнями звукового давления соответственно 84, 77 и 71 дБ. Требуется определить уровень громкости сложного звука. По кривым равного уровня громкости (рис. 2.4) определяем уровень громкости каждой составляющей. Они все попадают на одну кривую, поэтому у всех один уровень громкости 80 фон. По графику на рис. 2.5 определяем громкость каждой составляющей – 21,5 сон. Общая громкость равна сумме – 64,5 сон. По этому же графику находим общий уровень громкости – 93 фона.

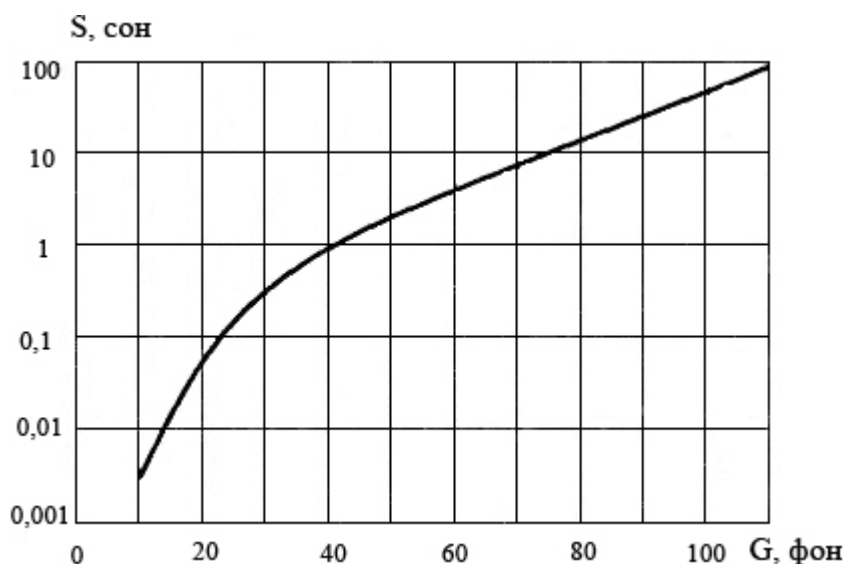


Рис. 2.5. Громкость тонального звука от его уровня громкости

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

**Таблица 2.1**  
**Средний уровень громкости и громкость наиболее часто встречающихся звуков и шумов**

Источник шума и место его измерения	Уровень громкости, фон	Громкость, сон
1	2	3
<b>Шумы вне помещений</b>		
Авиационный мотор на расстоянии 5 м	116–120	346–556

Окончание таблицы 2.1

1	2	3
Громкий автомобильный сигнал на расстоянии 8 м	95–100	57–88
Электропоезд на эстакаде на расстоянии 6 м	90	38
Шум в поезде метро во время движения	85–90	25–38
Автобус (полный ход) на расстоянии 5 м	85–88	25–32,2
Трамвай на расстоянии 10–20 м	80–85	17,1–25
Грузовой автомобиль в городе на расстоянии 5–20 м	60–75	4,35–11,4
Легковой автомобиль в городе на расстоянии 5–20 м	50–65	2,2–5,87
Свисток милиционера на расстоянии 20 м	70	7,95
Шумная улица без трамвайного движения	60–75	4,35–11,4
Обычный средний шум на улице	55–60	3,08–4,35
То же, в момент затишья днем	40	1,0
Тихая улица (без движения транспорта)	30–35	0,36–0,62
Тихий сад	20	0,097
<b>Производственные шумы</b>		
Удары молота по стали, клепальная машина на расстоянии 2–4 м	110–113	215–288
Котельные цехи	100–103	88–116
Общий шум в ткацком цехе	96–100	62–88
Деревообрабатывающая фабрика	96–98	62–74
<b>Театры, школы, больницы</b>		
Симфонический оркестр	80–100	17,1–88
Аплодисменты	60–75	4,4–11,4
Громкая музыка по радио	80	17,1
Радиоцентр (студия во время исполнения)	40–50	1–2,2
Актзовые залы в школах во время перерыва	55–62	3,1–4,7
Библиотеки	25–30	0,2–0,36
Больницы	20–30	0,1–0,36
Шумное собрание	65–70	5,87–7,95
Громкий разговор на расстоянии 1 м	65–70	5,87–7,95
Обычный на расстоянии 1 м	55–60	3,08–4,35
<b>Жилые помещения</b>		
Разговор трех человек в обычной комнате	45–50	1,5–2,2
Шепот средней громкости на расстоянии 1 м	20	0,1
<b>Степени музыкальной громкости</b>		
Форте-фортиссимо	100	88
Фортиссимо	90	38
Форте	80	17,1
Меццо-форте	70	7,95
Меццо-пиано	60	4,35
Пиано	50	2,2
Пианиссимо	40	1,0
Пиано-пианиссимо	30	0,36

### 2.6.3. Временные характеристики слуха

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

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

### 2.7. Адаптация слуха

При воздействии на барабанную перепонку уха достаточно длительного звука большой интенсивности воспринимаемая громкость постепенно уменьшается. Это значит, что во время действия длительного громкого звука падает чувствительность уха. После прекращения действия звука чувствительность слуха постепенно восстанавливается. Это явление называется адаптацией слуха. Явление адаптации слуха поясняется рис. 2.6.

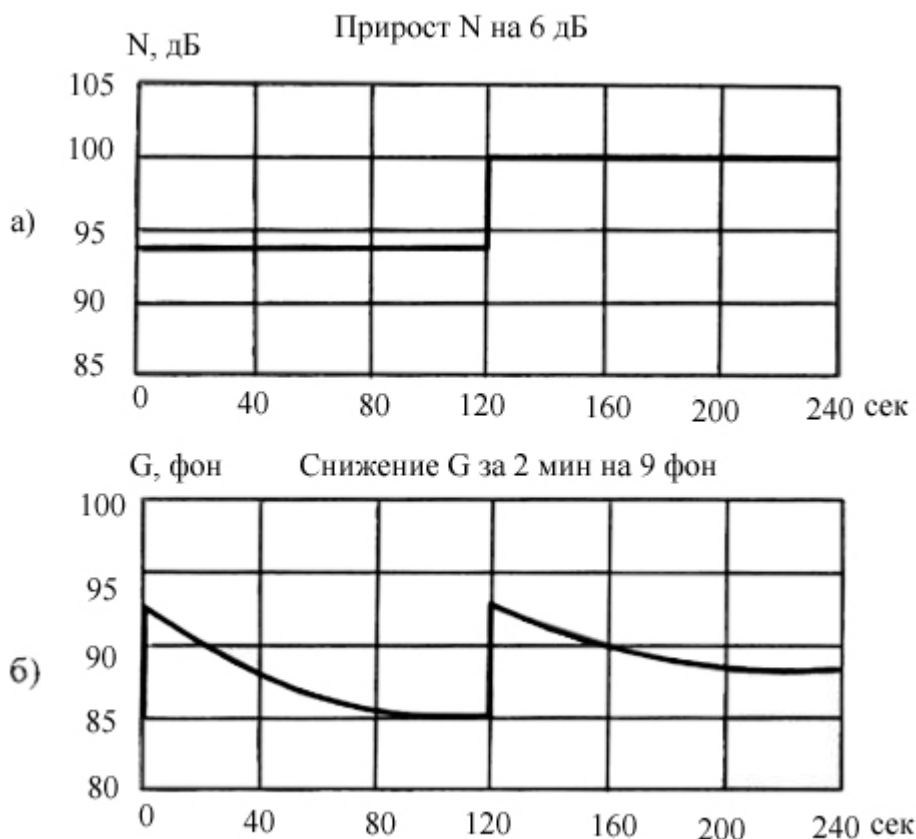


Рис. 2.6. Графики адаптации слуха при увеличении звукового давления

### 2.8. Маскировка во временной области

До сих пор мы определяли коэффициент маскировки в предположении, что маскируемый и маскирующий звуки присутствуют одновременно.

но. Однако из-за явления адаптации слуха возникают ситуации, когда достаточно громкие звуки маскируют, делают практически неслышимыми звуки, следующие за ними. В некоторых случаях маскируются предшествующие звуки.

Такой вид маскировки, когда звуки не перекрываются по времени, называется временной маскировкой. Она разделяется на предмаскировку и послемаскировку. Послемаскировка проявляется на интервале времени 100–200 мс после окончания маскирующего звука. Предмаскировка проявляется на значительно более короткие временные интервалы около 10 мс.

## 2.9. Бинауральный слух

**Бинауральным слухом** называется его способность определять направление прихода звуковой волны, т. е. локализовать положение источника звука в пространстве. Эта способность достигается благодаря пространственной несовместности двух ушей в сочетании с экранирующим влиянием головы. Это приводит к тому, что всегда имеет место неидентичность возбуждения правого и левого уха. Этот факт обеспечивает человеку возможность воспринимать пространственный звуковой мир и оценивать перемещение источников звука в пространстве.

К числу основных свойств бинаурального слуха можно отнести:

- пространственную локализацию;
- эффект предшествования;
- бинауральное суммирование громкости;
- бинауральную демаскировку.

Такие возможности слуха достигаются благодаря трем факторам.

**Временным** – возникающим из-за несовпадения моментов воздействия одинаковых фаз звука на левое и правое ухо.

**Амплитудным** – возникающим из-за неодинаковой величины звуковых давлений на левое и правое ухо вследствие дифракции звуковой волны вокруг головы, например, образования акустической тени со стороны, обратной источнику звука.

**Спектральным** – возникающим из-за разницы в спектральном составе звуков, воспринимаемых левым и правым ухом, вследствие неодинакового экранирующего влияния головы и ушных раковин на низких и высоких звуковых частотах.

### Вопросы по лекции

1. Какая наука занимается изучением слуха и восприятием звука?
2. Что позволяют разрабатывать результаты, полученные в психоакустике?
3. Что требуется для повышения качества воспроизведения звука?

4. Что такое субполосное кодирование?
5. Чем определяются требования к телефонам, микрофонам, громкоговорителям, к аппаратуре записи и воспроизведения звука?
6. Какова зависимость высоты тона от частоты сигнала?
7. В каких единицах измеряется высота тона?
8. Дайте определение порога слышимости.
9. Какие частоты называются инфразвуковыми, а какие ультразвуковыми?
10. Что называется порогом болевого ощущения?
11. Что гласит психофизический закон, установленный Вебером в 1846 году?
12. Что называется маскировкой?
13. Как рассчитывается коэффициент маскировки  $K_m$ ?
14. Единица измерения уровня громкости.
15. Что называется громкостью?
16. В каких единицах измеряется громкость?
17. Что называется бинауральным слухом?

## Лекция 3. РАБОТА СО ЗВУКОМ.

### ЗВУКОВЫЕ СИГНАЛЫ

#### 3.1. Определения

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

При оценке трактов вещания и связи полагают, что каждый акустический сигнал почти всегда является случайным в вероятностном смысле и несет в себе информацию, соответствующую его объему. Строго говоря, звуковые сигналы художественных программ, не могут рассматриваться как случайные в вероятностном смысле, так как они могут быть хорошо известны слушателям, и поэтому не нести в себе информацию. Для слушателя эти сигналы служат средством удовлетворения эстетических потребностей, а не получения информации. Однако при расчете каналов звукопередачи эти сигналы считают несущими информацию, соответствующую их объему. Хотя в музыкальных сигналах очень большие участки могут иметь периодический характер, в среднем, для больших интервалов времени их также можно рассматривать как случайные. Поэтому параметры звуковых сигналов определяются их распределениями по уровню, по частотному диапазону и во времени<sup>1</sup>.

Ко вторичным звуковым сигналам относятся сигналы, воспроизводимые электроакустическими устройствами, то есть первичные сигналы, прошедшие по электроакустическим трактам связи и вещания и соответственно видоизмененные по своим параметрам.

#### 3.2. Динамический диапазон

В процессе любой передачи уровень акустического сигнала непрерывно изменяется, причем диапазон его изменения может быть довольно широким. На рис. 3.1, а показана зависимость уровня сигнала от времени, называемая **уровнеграммой**. Обычно ее приводят для уровня, определенного при постоянной времени измерителя, равной или 150–200 мс (субъективная уровнеграмма), или 20–30 мс (объективная уровнеграмма).

---

<sup>1</sup> Определению параметров звуковых сигналов посвящены многочисленные исследования. В частности, имеются данные по распределению плотности вероятности мгновенных значений сигнала, распределению выбросов и пауз, определены текущая, среднeminутная и долговременная мощности сигнала и др.

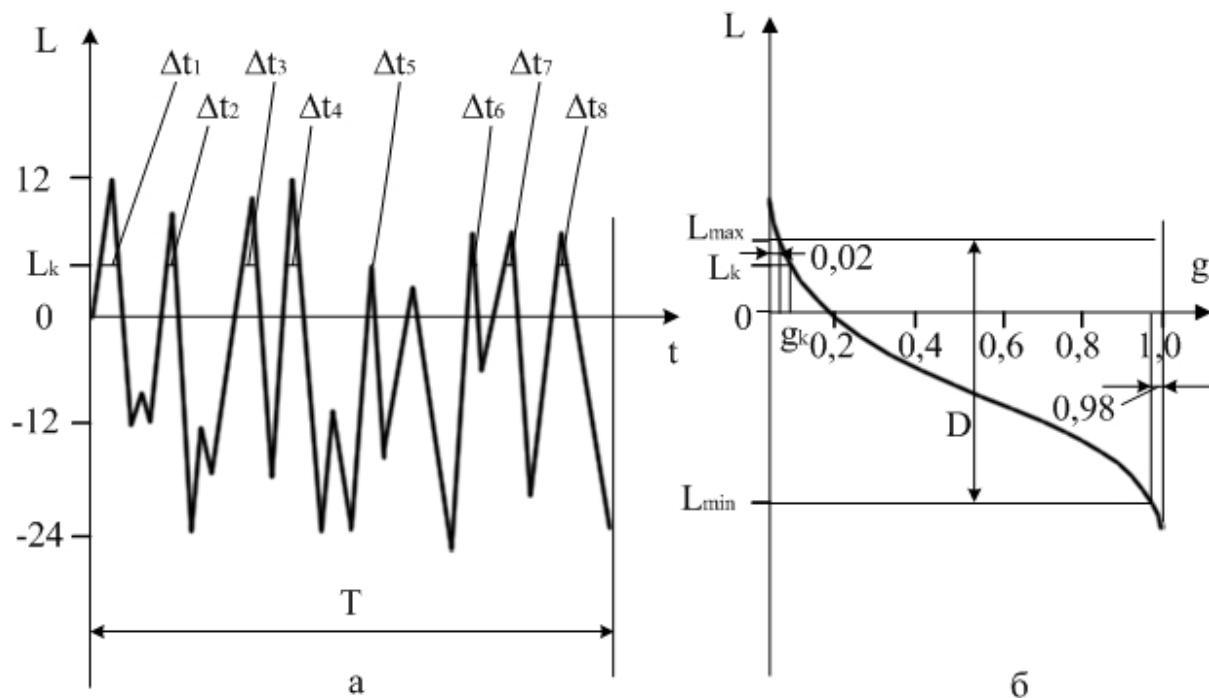


Рис. 3.1. К определению динамического диапазона:  
 а) уровнеграмма; б) построение по ней интегрального распределения

Динамический диапазон сигнала всегда необходимо сопоставлять с динамическим диапазоном канала звукопередачи  $D_K$ :

$$D_K = 20 \lg \frac{U_{НОМ}}{U_{Ш}} - (\Delta N_1 + \Delta N_2),$$

где  $U_{Ш}$  – уровень шума в канале;

$U_{НОМ}$  – номинальное напряжение;

$\Delta N_1$  – уровень перекрытия помех и шумов, дБ (не менее 10 дБ);

$\Delta N_2$  – допуск на перегрузку (3–6) дБ.

Динамический диапазон для некоторых видов звуковых сигналов приведен в таблице 3.1.

Таблица 3.1

**Динамический диапазон для некоторых видов звуковых сигналов**

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

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

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

### 3.3. Средний уровень

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

### 3.4. Частотный диапазон и спектры

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

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

Спектральная плотность  $G = I_{\Delta f} / \Delta f$ , где  $I_{\Delta f}$  – интенсивность, измеренная с помощью узкополосных фильтров.

Для удобства оценки введена логарифмическая мера плотности спектра аналогично уровню интенсивности. Эту меру называют уровнем спектральной плотности или **спектральным уровнем**. Спектральный уровень  $B = 10 \lg I / I_0$ , где  $I_0 = 10^{-12}$  Вт/м<sup>2</sup> – интенсивность, соответствующая нулевому уровню.



### 3.5. Временные характеристики акустического сигнала

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

### 3.6. Первичный речевой сигнал

Каждому человеку свойственна своя манера произносить звуки речи (своего рода устный почерк). Произношение звуков речи зависит, например, от ударения, соседних звуков и других факторов. Физические реализации ограниченного числа обобщенных звуков речи называются **фонемами**. Фонема – это то, что человек хочет произнести, а звук речи – это то, что человек фактически произносит. Фонема по отношению к звуку речи играет ту же роль, что и образцовая буква, называемая **графемой** (например, печатная курсивом) по отношению к ее рукописной форме в конкретном написании.

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

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

тембр (окраска) голоса могут служить для опознавания человека. При этом степень достоверности опознавания выше, чем по отпечаткам пальцев. Это свойство используют в разработанной в последнее время аппаратуре, реагирующей только на определенные голоса.

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

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

Резкое увеличение объема речевого сигнала объясняется следующим. Речевой сигнал представляет собой спектрально-модулированную несущую  $\zeta = E(\omega)F(\omega)$ , где  $E(\omega)$  – спектр импульсов основного тона или шума;  $F(\omega)$  – фильтровая (модуляционная) функция речевого тракта.

В результате спектральной модуляции изменяется соотношение между частотными составляющими несущей, то есть изменяется форма огибающей ее спектра (появляются **форманты** и **антиформанты**). Почти вся информация о звуках речи заключается в этой спектральной огибающей и ее временном изменении. Эти изменения происходят медленно (в темпе произнесения звуков), поэтому передача сведений об огибающей и ее изменении не требует пропускной способности тракта более 100 бит/с. Но для передачи широкополосной несущей с ее широким динамическим диапазоном требуется очень большая пропускная способность. Кроме того, речевой сигнал при образовании в речевом тракте приобретает много информации, не относящейся к смыслу передаваемой речи (например, фазовую информацию). Эта информация называется сопутствующей. Для ее передачи также расходуется пропускная способность тракта. Из этого следует, что избыточность речевого сигнала лишь немного превышает избыточность телеграфного сигнала с таким же сообщением (речевой сигнал отличается от телеграфного лишь информацией об эмоциях и личности говорящего). Для передачи смысла достаточно передавать сведения о форме огибающей спектра речи, а также об изменении основного тона речи и переходах тон-шум.

### 3.7. Вторичный сигнал

В идеальном случае вторичный сигнал должен точно воспроизводить первичный, но это не всегда требуется, так как слух человека может и не заметить их несоответствие. К тому же на практике их точное соответствие часто невозможно или очень трудно осуществить. При художественном вещании, телевидении и звукозаписи надо стремиться к этому соответствию в пределах, при которых слуховое ощущение, создающееся у слушателя, было бы близко к тому ощущению, которое он получает, находясь в месте исполнения данной программы при создании достаточно хороших акустических условий. Для информационных программ вещания и телефонной связи этого соответствия добиваются в первую очередь для получения полной понятности речи, а затем для достаточно высокого качества звучания. Только в этом случае необходимо стремиться к более точному соответствию вторичного сигнала первичному. В обоих случаях существенную роль играют экономические соображения.

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

#### Вопросы по лекции

1. Приведите примеры динамических диапазонов для некоторых видов звуковых сигналов.
2. Как можно определить средний уровень интенсивности акустического сигнала?
3. Объясните различия первичных и вторичных звуковых сигналов.
4. Как можно охарактеризовать понятие тембра?
5. Что называется периодом основного тона голоса  $T_0$  и частотой основного тона?
6. Как ведется борьба с шумами и помехами?
7. Что такое дискретизация?

## Лекция 4. РАБОТА СО ЗВУКОМ.

### ЦИФРОВОЕ ПРЕДСТАВЛЕНИЕ ЗВУКОВЫХ СИГНАЛОВ

#### 4.1. Аналого-цифровое преобразование

Преобразование аналогового звукового сигнала в цифровой включает в себя несколько этапов. Сначала аналоговый звуковой сигнал подается на аналоговый фильтр, который ограничивает полосу частот сигнала и устраняет помехи и шумы. Затем из аналогового сигнала с помощью схемы выборки/хранения выделяются отсчеты: с определенной периодичностью осуществляется запоминание мгновенного уровня аналогового сигнала. Далее отсчеты поступают в аналого-цифровой преобразователь (АЦП), который преобразует мгновенное значение каждого отсчета в цифровой код или числа. Полученная последовательность бит цифрового кода, собственно, и является звуковым сигналом в цифровой форме. В результате преобразования непрерывный аналоговый звуковой сигнал превращается в цифровой – дискретный как по времени, так и по величине. Для примера на рис. 4.1 показана структурная схема канала цифровой записи звука.

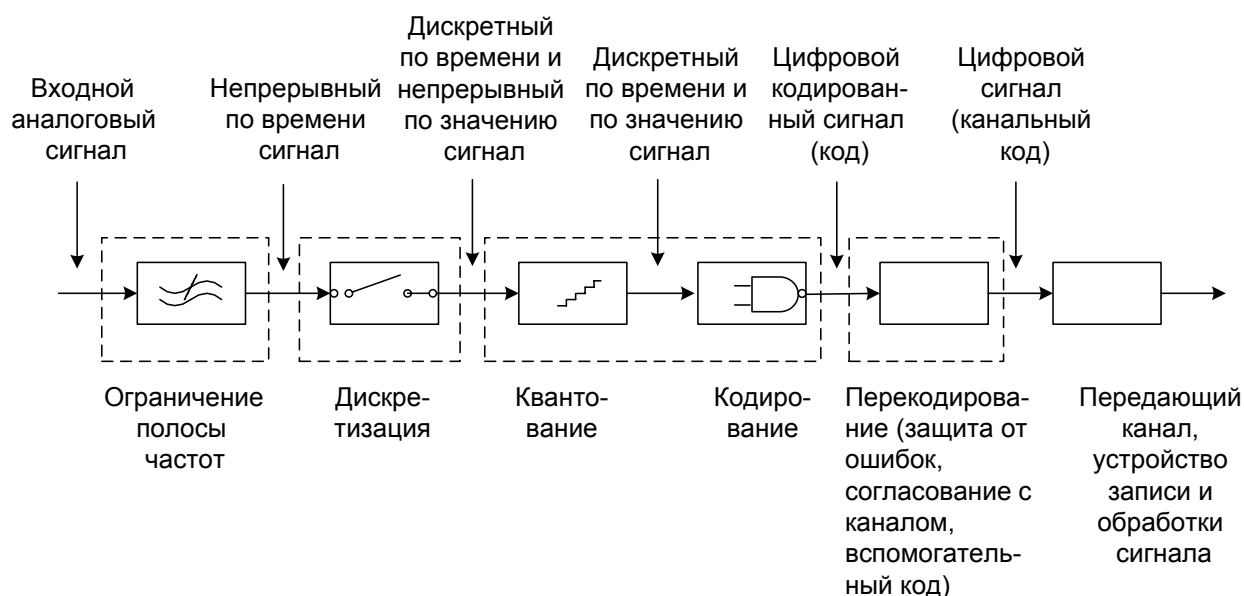


Рис. 4.1. Обобщенная структурная схема канала цифровой записи звука

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

### 4.1.1. Дискретизация

Важнейшим этапом аналого-цифрового преобразования является дискретизация аналогового сигнала. Вместо термина «дискретизация» в технической литературе иногда употребляют термин «выборка», а в литературе, посвященной обработке звука используется понятие – «сэмплирование»<sup>2</sup>.

По определению, дискретизация – это процесс взятия отсчетов непрерывного во времени сигнала в равноотстоящих друг от друга по времени точках. Иными словами, в процессе дискретизации измеряется и запоминается уровень аналогового сигнала. Через заданный интервал времени, который называется интервалом дискретизации, процедура повторяется. Для качественного преобразования аналогового сигнала в цифровой необходимо производить достаточно большое количество отсчетов даже в течение одного периода изменения аналогового сигнала, другими словами, значение частоты дискретизации не может быть произвольным.

И действительно, значение частоты дискретизации фактически определяет ширину полосы частот сигнала, который может быть записан с помощью используемой цифровой системы. Ширина этой полосы не может быть больше половины значения частоты дискретизации, как определяет **теорема отсчетов (Котельникова-Найквиста)**. Эта теорема имеет важнейшее значение в технике записи и передачи звука в цифровой форме. Теорема гласит: **сигнал, спектр частот которого занимает область от  $F_{МИН}$  до  $F_{МАКС}$  (низкочастотный звуковой сигнал), может быть полностью представлен своими дискретными отсчетами с интервалом  $T_д$ , если  $T_д$  не превышает  $1/2F_{МАКС}$ . Другими словами, частота дискретизации  $f_д = 1/T_д$  в процессе преобразования должна быть, как минимум, вдвое больше наивысшей частоты звукового сигнала  $F_{МАКС}$ .**

Почему же именно в два раза? Да потому, что спектр сигнала, преобразованного с помощью АЦП в цифровую форму, имеет периодический характер.

В соответствии с теоремой Фурье сигнал любой формы может быть представлен в виде суммы простейших синусоидальных колебаний разной частоты и амплитуды. По окончании аналого-цифрового преобразования

---

<sup>2</sup> С английского языка слово Sample дословно переводится как «образец». Поэтому это слово в мультимедийной и профессиональной терминологии имеет несколько значений для обозначения разных типов «образцов». Чаще всего сэмплом называют промежуток времени между двумя измерениями аналогового сигнала. Кроме промежутка времени сэмплом называют последовательность цифровых данных, полученных в результате аналого-цифрового преобразования, а сам процесс преобразования – сэмплированием. В иностранных журнальных публикациях такой термин как частота дискретизации вы не встретите, но в изобилии столкнетесь с частотой сэмплирования, хотя эти термины обозначают одно и то же. Термин «дискретизация» нам более привычен. Поэтому далее мы будем использовать термин «дискретизация».

звуковой сигнал, представленный в цифровой форме, содержит, кроме низкочастотных, соответствующих исходному аналоговому сигналу, еще и высокочастотные компоненты (рис. 4.2). Эти компоненты есть повторение низкочастотного спектра сигнала в виде боковых полос с центрами в точках, кратных частоте дискретизации ( $f_d$ ,  $2f_d$ ,  $3f_d$ ,  $4f_d$  и т. д.).



Рис. 4.2. Перекрывание спектров сигнала при дискретизации

Если уменьшить частоту дискретизации, то произойдет наложение (перекрывание) низкочастотной части спектра и боковой полосы с центром в точке. Наложение спектров приведет к появлению новых спектральных составляющих в сигнале и, следовательно, к невозможности его правильного восстановления.

Классическим примером наложения спектров является случай, когда при просмотре кинофильма кажется, что колесо движущейся кареты крутится со скоростью, не соответствующей скорости движения кареты, или даже в обратную сторону. Возникновение этого эффекта обусловлено тем, что скорость смены кадров (частота дискретизации изображения) мала по сравнению с угловой скоростью вращения колеса.

Чтобы при записи звукового сигнала избежать наложения спектров, перед АЦП устанавливается фильтр низких частот (ФНЧ), подавляющий все частоты, лежащие выше частоты дискретизации. При этом желательно, чтобы фронты АЧХ этого фильтра были как можно круче.

Если учесть, что человек способен слышать звуковые колебания, частота которых находится в диапазоне от 16–20 Гц до 20 кГц, и с позиций

теоремы отсчетов взглянуть на требования к частотным характеристикам высококачественной звукотехники (например, проигрывателей аудиокомпакт-дисков), становится ясно, что частота дискретизации исходного звукового сигнала должна составлять не менее 40 кГц. Реально для подобных систем частота дискретизации выбирается не менее 44,1 кГц. Стандартное значение частоты дискретизации большинства звуковых карт составляет 44,1 и 48,0 кГц.

Итак, результатом дискретизации является дискретный во времени сигнал, представляющий собой последовательность отсчетов – мгновенных значений уровня аналогового сигнала. Чем выше частота дискретизации, тем более точно будет восстановлен звуковой сигнал.

Процедура дискретизации технически реализуется с помощью устройства выборки/хранения. В качестве запоминающего элемента обычно используют конденсатор, заряжающийся до уровня напряжения входного сигнала. При этом потенциал заряда конденсатора соответствует мгновенному значению напряжения сигнала. Напряжение на конденсаторе сохраняется неизменным в течение некоторого отрезка времени, называемого временем хранения. В идеальном случае взятие отсчета должно происходить мгновенно, реально же длительность этого процесса составляет приблизительно 1 мкс.

#### 4.1.2. Квантование

После дискретизации происходит второй этап аналого-цифрового преобразования – квантование отсчетов. В процессе квантования производится измерение мгновенных значений уровня сигнала, полученных в каждом отсчете, причем осуществляется оно с точностью, которая напрямую зависит от количества разрядов, используемых для записи значения уровня.

Если, задав длину  $N$ -кодového слова, записать значения уровня сигнала с помощью двоичных чисел, то количество возможных значений будет равно  $2^N$ . Столько же может быть и уровней квантования. Например, если значение амплитуды отсчета представляется 16-разрядным кодовым словом, то максимальное количество градаций уровня сигнала (уровней квантования) будет равно 65536 ( $2^{16}$ ). При 8-разрядном представлении будем иметь 256 ( $2^8$ ) градаций уровня.

**Шумы квантования.** Преобразование сигнала из аналоговой формы в цифровую можно произвести только с какой-то степенью точности, при этом, чем выше частота дискретизации и разрядность АЦП, тем точнее происходит преобразование.

Искажения сигнала, возникающие в процессе квантования отсчетов, влекут за собой потерю информации, которую при обратном цифро-аналоговом преобразовании в ходе воспроизведения записанного звукового сигнала ликвидировать или уменьшить практически невозможно.

Например, если преобразовать линейно нарастающий по напряжению аналоговый сигнал, то дискретный и восстановленный сигналы будут различаться на величину напряжения ошибки  $U_{ош}$  (рис. 4.3). При записи звука зависимость ошибки от времени гораздо сложнее, а ее спектр подобен спектру белого шума и занимает частотный диапазон от нуля до частоты дискретизации. Появление ошибок квантования при записи звукового сигнала в цифровой форме эквивалентно добавлению к восстановленному сигналу некоторого шума. Поэтому ошибки квантования называются шумом квантования. Шум квантования можно рассматривать и как специфические искажения сигнала, особенно заметные при малых его уровнях. Уровень шума квантования обычно измеряется в присутствии сигнала как уровень (в децибелах) относительно максимального значения сигнала. Чем меньше этот уровень, тем выше качество звука. Достижимый уровень шума определяется разрядностью квантования и частотой дискретизации.

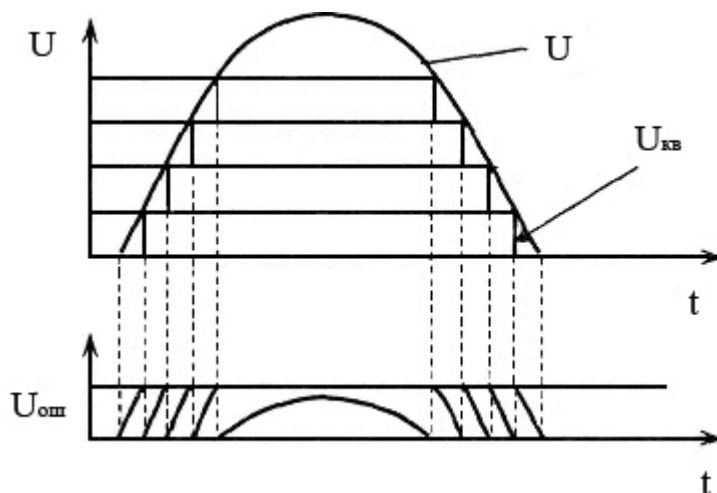


Рис. 4.3. Ошибки квантования

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

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



### 4.1.3. Передискретизация (оверсэмплинг)

Для того чтобы осуществить аналого-цифровое преобразование с высоким качеством, необходимо выполнить ряд условий.

Прежде всего, при оцифровке звукового сигнала следует использовать как можно более высокую частоту дискретизации: чем выше будет частота дискретизации, тем более качественно будет восстановлен исходный сигнал. К сожалению, пропорционально увеличению частоты дискретизации возрастает поток цифровых данных в канале звукозаписи, а также объем памяти, необходимой для хранения звукового сигнала в цифровой форме.

Другое условие аналого-цифрового преобразования заключается в том, что перед дискретизацией необходимо ограничить спектр входного сигнала с помощью фильтра низкой частоты (ФНЧ). Он должен удалить все гармоники с частотами, лежащими выше частоты дискретизации, и тем самым предотвратить наложение спектров.

В современных АЦП проблема фильтрации с целью устранения высокочастотных компонент спектра решается с помощью передискретизации – дискретизации на повышенной частоте. Термину передискретизация в зарубежной технической литературе соответствует термин оверсэмплинг, который в дальнейшем и будем использовать.

При оверсэмплинге диапазон частот входного аналогового звукового сигнала ограничивается с помощью ФНЧ низкого порядка (обычно 3–5-го), имеющего линейную фазовую характеристику и практически не искажающего импульсный сигнал. Частота среза фильтра выбирается значительно выше частоты полезного сигнала и составляет 25–30 кГц. В результате исключаются фазовые искажения, характерные для аналоговых фильтров высокого порядка, и подавление полезных сигналов высших частот. Отфильтрованный сигнал, имеющий ограниченный по частоте спектр, подвергается дискретизации на повышенной частоте, что исключает наложение и искажение спектра.

Далее дискретные отсчеты сигнала преобразуются в последовательность чисел с помощью АЦП, причем поток цифровых данных включает и нежелательные высокочастотные компоненты спектра.

Полученные цифровые данные подвергаются цифровой фильтрации. Цифровой фильтр высокого порядка с крутым срезом частотной характеристики изготовить гораздо проще. Причем, благодаря линейности фазовой характеристики цифрового фильтра, фазовые искажения сигнала будут отсутствовать. После цифрового фильтра сигнал будет иметь спектр, корректно ограниченный по частоте.

После цифровой фильтрации частота дискретизации сигнала понижается до удвоенного значения наивысшей полезной частотной составляющей путем удаления «избыточных» отсчетов.

В результате оверсэмплинга нежелательные высокочастотные составляющие будут ликвидированы, в то время как высокочастотные составляющие исходного звукового сигнала будут сохранены.

## 4.2. Цифро-аналоговое преобразование

Для воспроизведения звукового сигнала, записанного в цифровой форме, необходимо преобразовать его в аналоговую форму, то есть осуществить цифро-аналоговое преобразование сигнала (рис. 4.4).

Цифро-аналоговое преобразование производится в два этапа.

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

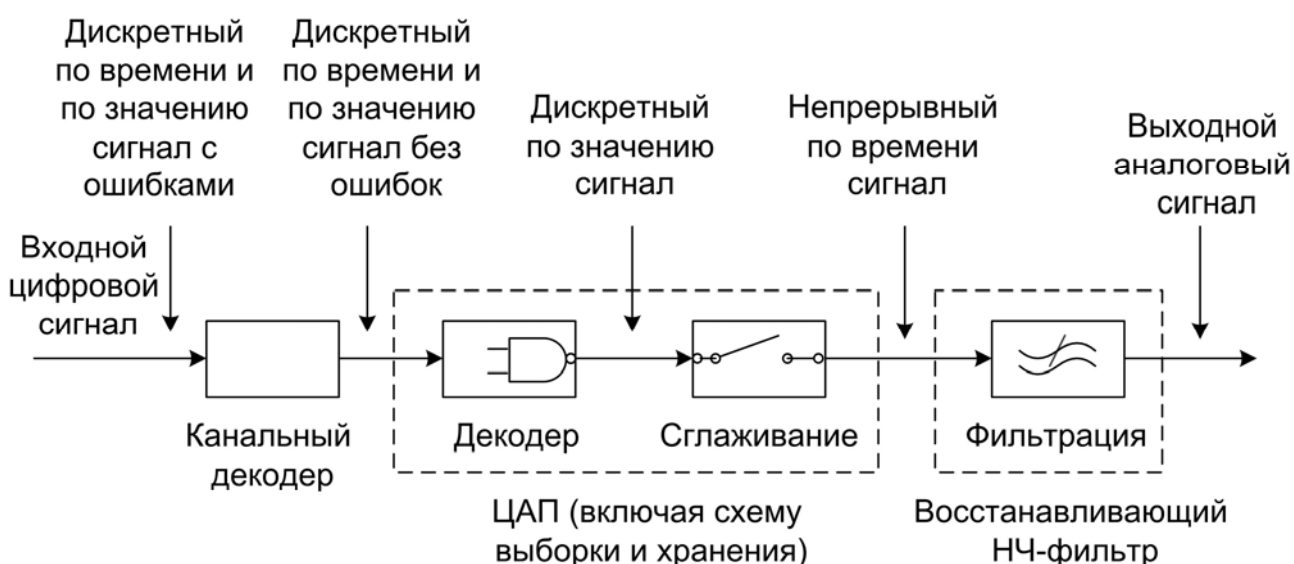


Рис. 4.4. Обобщенная схема преобразования цифрового сигнала в аналоговый

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

Сразу после первого этапа цифро-аналогового преобразования сигнал представляет собой серию узких импульсов, имеющих многочисленные высокочастотные спектральные компоненты. На аналоговый фильтр в этом случае возлагается задача полностью пропустить сигнал нужного частотного диапазона (например, 0–24 кГц) и как можно сильнее подавить ненужные высокочастотные компоненты. Аналоговому фильтру выполнить такие противоречивые требования не под силу.

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

### **Вопросы по лекции**

1. О чем гласит теорема отсчетов Котельникова-Найквиста?
2. Каково стандартное значение частоты дискретизации большинства звуковых карт?
3. Как повысить качество цифрового звука и увеличить эффективную разрядность АЦП?
4. Чему равна скорость цифрового потока  $V$  при передаче одного сигнала, если  $f = 48$  кГц и  $\Delta A = 16$  бит/отсчет? И какой суммарной пропускной способности канала связи требуется при передаче звукового сигнала форматов 5.1 (Dolby Digital)?
5. Сколько бит/с способен человек сознательно обрабатывать своими органами чувств?
6. Что происходит на этапах цифро-аналогового преобразования?

### **Библиографический список**

1. Катунин, Г. П. Основы мультимедиа. Звук и видео / Г. П. Катунин : монография. – Новосибирск, СибГУТИ, 2006. – 389 с.
2. Чепмен, Найджел. Цифровые технологии мультимедиа / Найджел Чепмен, Дженни Чепмен. 2-е изд. – М. : Диалектика, 2005. – 624 стр., с ил.
3. Андерсен, Бент Б. Мультимедиа в образовании / Бент Б. Андерсен, Катя ван ден Бринк – М. : Дрофа, 2007. – 224 с.