



§ 16

Кодирование звуковой информации

16.1. Звук и его характеристики

Звук — это распространяющиеся в воздухе, воде или другой среде волны с непрерывно меняющейся амплитудой и частотой (рис. 3.12).

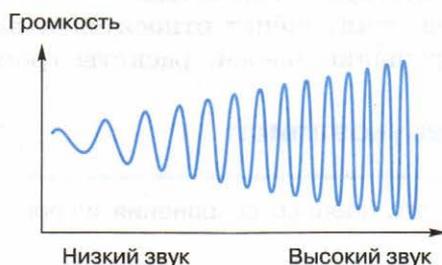


Рис. 3.12. Звуковая волна

Амплитуду звуковых колебаний называют звуковым давлением или силой звука. Эта величина характеризует воспринимаемую громкость звука. Абсолютную величину звукового давления измеряют в единицах давления — паскалях (Па). Самые слабые, едва различимые звуки имеют амплитуду около 20 мкПа ($2 \cdot 10^{-5}$ Па, так называемый порог слышимости). Самые сильные звуки, не



выводящие слуховые органы из строя, могут иметь амплитуду до 200 Па (так называемый болевой порог).

На практике вместо абсолютной используют относительную силу (уровень) звука, измеряемую в децибелах (дБ). Вот некоторые значения уровня звука:

Порог слышимости	0 дБ
Шелест листья, дыхание	10–20 дБ
Шёпот	20–30 дБ
Разговор средней громкости	50–60 дБ
Шумная улица	70–80 дБ
Громкая музыка	100–120 дБ
Самолёт на взлёте	120–130 дБ
Болевой порог	140 дБ

Частота определяется как количество колебаний в секунду и выражается в герцах (Гц). Чем больше частота, тем выше звук, и наоборот. Человек способен слышать звук в широком частотном диапазоне, но важное для жизни значение имеют только звуки от 125 до 8000 Гц. Например, звуковые волны в диапазоне 500–4000 Гц соответствуют человеческому голосу. Звучание детского голоса, пение птиц, шёпот относятся к высоким частотам. Звук контрабаса, рычание зверей, раскаты грома — к низким.

16.2. Понятие звукозаписи

Звукозапись — это процесс сохранения информации о параметрах звуковых волн.



Способы записи звука разделяются на аналоговые и цифровые. При аналоговой записи на носителе размещается непрерывный «слепок» звуковой волны. Так, на грампластинке пропечатывается непрерывная канавка, изгибы которой повторяют амплитуду и частоту звука.

16.3. Оцифровка звука

Чтобы компьютер мог обрабатывать звук, непрерывный звуковой сигнал должен быть преобразован в цифровую дискретную форму. Для этого его подвергают временной дискретизации и квантованию: параметры звукового сигнала измеряются не непрерывно, а через определённые промежутки времени (временная дискретизация); результаты измерений записываются в цифровом виде с ограниченной точностью (квантование).

Вообще говоря, в компьютер приходит не сам звук, а электрический сигнал, снимаемый с какого-либо устройства: например, микрофон преобразует звуковое давление в электрические колебания, которые в дальнейшем и обрабатываются.

Если записывается стереозвук (ведётся двухканальная запись), то оцифровке подвергается не один электрический сигнал, а сразу два и, следовательно, количество сохраняемой цифровой информации удваивается.

Сущность временной дискретизации заключается в том, что аналоговый звуковой сигнал разбивается на отдельные маленькие временные участки и для каждого такого участка устанавливается определённая величина интенсивности звука (рис. 3.13). Другими словами, через какие-то промежутки времени мы измеряем уровень аналогового сигнала. Количество таких измерений за одну секунду называется частотой дискретизации.

Частота дискретизации — это количество измерений громкости звука за одну секунду.

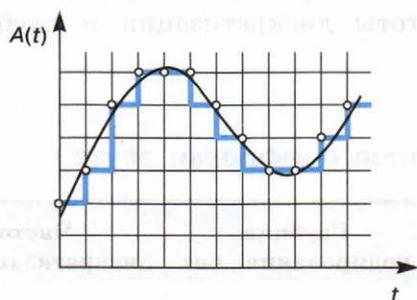


Рис. 3.13. Временная дискретизация звукового сигнала ($A(t)$ — амплитуда, t — время)

Частота дискретизации измеряется в герцах (Гц) и килогерцах (кГц). $1 \text{ кГц} = 1000 \text{ Гц}$. Частота дискретизации, равная 100 Гц, означает, что за одну секунду проводилось 100 измерений громкости звука.

Качество звукозаписи зависит не только от частоты дискретизации, но также и от глубины кодирования звука.

Глубина кодирования звука или разрешение — это количество информации, которое необходимо для кодирования дискретных уровней громкости цифрового звука.

В результате измерений звукового сигнала (см. рис. 3.13) на каждой его «ступеньке» будет получено некоторое значение громкости, при этом все результаты измерений будут лежать в некотором диапазоне.

Пусть под запись одного результата измерения громкости в памяти компьютера отведено n бит. Вы знаете, что это позволяет закодировать ровно 2^n разных результатов измерений. Так, при $n = 8$ можно закодировать 256 разных результатов измерений громкости звука. Поэтому весь диапазон, в котором могут находиться результаты измерений громкости звука, можно разбить на 256 разных поддиапазонов — уровней громкости звука, каждому из которых присвоить свой уникальный код. После этого каждый имеющийся результат измерений громкости звука можно соотнести с некоторым поддиапазоном, в который он попадает, и кодировать его номером (кодом) соответствующего уровня громкости.

В зависимости от ситуации на практике используются разные значения частоты дискретизации и глубины кодирования (табл. 3.13).

Таблица 3.13

Примеры параметров оцифровки звука

Название	Глубина кодирования, бит	Частота дискретизации, кГц	Число каналов
Телефонная связь	8	8	1
CD	16	44,1	2

Пример. Оценим объём звукового стереоаудиофайла с глубиной кодирования 16 бит и частотой дискретизации 44,1 кГц, который хранит звуковой фрагмент длительностью звучания 15 секунд.

Объём такого звукового фрагмента равен:

$$\begin{aligned} 2 \text{ (канала)} \cdot 16 \text{ бит} \cdot 44\,100 \text{ Гц} \cdot 15 \text{ с} = \\ = 2\,646\,000 \text{ байт} \approx 2\,584 \text{ Кбайта.} \end{aligned}$$

Увеличивая частоту дискретизации и глубину кодирования, можно более точно сохранить и впоследствии восстановить форму звукового сигнала. При этом объём сохраняемых данных будет увеличиваться.

Важно понимать, каких параметров оцифровки достаточно, чтобы сохраняемый звук был достаточно близок к исходному, а содержащий его файл имел минимально возможный объём. В начале 30-х годов прошлого века было установлено, что это возможно, если частота временной дискретизации будет в два раза выше максимальной частоты измеряемого сигнала.

В 1928 году американский учёный Гарри Найквист высказал утверждение, что частота дискретизации должна быть в два или более раза выше максимальной частоты измеряемого сигнала. В 1933 году наш соотечественник В. А. Котельников и независимо от него американец Клод Шеннон в 1949 году сформулировали и доказали теорему, более сильную чем утверждение Найквиста, о том, при каких условиях и как по дискретным значениям можно восстановить форму непрерывного сигнала.

САМОЕ ГЛАВНОЕ

Звук — это распространяющиеся в воздухе, воде или другой среде волны с непрерывно меняющейся амплитудой и частотой.

Чтобы компьютер мог обрабатывать звук, непрерывный звуковой сигнал должен быть преобразован в цифровую дискретную форму. Для этого его подвергают временной дискретизации и квантованию: параметры звукового сигнала измеряются не непрерывно, а через определённые промежутки времени (временная дискретизация); результаты измерений записываются в цифровом виде с ограниченной точностью (квантование).

Таким образом, при оцифровке звука искажение сохраняемого сигнала происходит дважды: во-первых, при дискретизации теряется информация об истинном изменении звука между изме-

рениями, а во-вторых, при квантовании сохраняются не точные, а близкие к ним дискретные значения.

Объём оцифрованного звукового фрагмента в битах находится как произведение частоты дискретизации в Гц, глубины кодирования звука в битах, длительности звучания записи в секундах и количества каналов.



Вопросы и задания

1. Каким образом происходит преобразование непрерывного звукового сигнала в дискретный цифровой код?
2. Как частота дискретизации и глубина кодирования влияют на качество цифрового звука?
3. Производится четырёхканальная (квадро) звукозапись с частотой дискретизации 32 кГц и 32-битным разрешением. Запись длится 4 минуты, её результаты заносятся в файл, сжатие данных не производится. Определите приблизительно размер полученного файла (в мегабайтах). В качестве ответа укажите ближайшее к размеру файла целое число, кратное 10.
4. Музыкальный фрагмент был записан в формате моно, оцифрован и сохранён в виде файла без использования сжатия данных. Размер полученного файла — 49 Мбайт. Затем тот же музыкальный фрагмент был записан повторно в формате стерео (двухканальная запись) и оцифрован с разрешением в 4 раза выше и частотой дискретизации в 3,5 раза меньше, чем в первый раз. Сжатие данных не производилось. Укажите в мегабайтах размер файла, полученного при повторной записи.
5. Музыкальный фрагмент был оцифрован и записан в виде файла без использования сжатия данных. Получившийся файл был передан в город А по каналу связи за 32 секунды. Затем тот же музыкальный фрагмент был оцифрован повторно с разрешением в 3 раза выше и частотой дискретизации в 3 раза выше, чем в первый раз. Сжатие данных не производилось. Полученный файл был передан в город Б. Пропускная способность канала связи с городом Б в 2 раза выше, чем канала связи с городом А. Сколько секунд длилась передача файла в город Б?